

Gateisolatoren für MOS-Feldeffekttransistoren

DISTRIBUTION STATEMENT A:
Approved for Public Release -
Distribution Unlimited

Dissertation
von
Thomas Pompl

2000

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE 2000	3. REPORT TYPE AND DATES COVERED Dissertation	
4. TITLE AND SUBTITLE Gateisolatoren fuer MOS-Feldeffekttransistoren Gate Isolators for MOSFETs		5. FUNDING NUMBERS	
6. AUTHOR(S) Thomas Pompl		5. FUNDING NUMBERS	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Fakultaet fuer Elektrotechnik, Insitut fuer Physik		8. PERFORMING ORGANIZATION Report Number REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Universitaet der Bundeswehr Muenchen		10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES Text in German. Title and abstract in German and English.			
12a. DISTRIBUTION/AVAILABILITY STATEMENT Distribution A: Public Release.		12b. DISTRIBUTION CODE	
ABSTRACT (Maximum 200 words) The dissertation deals with Metal Oxide Semiconductor (MOS) field-effect transistors and how gate isolators have come to play a role in these semiconductors since the latter's introduction in the early 1970s. Testing the reliability of MOSFETs under hot carrier and bias temperature stress, the author then proceeds to the nitrogen barrier against boron diffusion. The bulk of the work concentrates on the tunneling of charged particles through a thin silicon dioxide layer as it functions in dual work function complementary MOS technology. Various gate leak flow mechanisms are used in examples with alternated MOSes. An analysis of the reliability with silicon dioxide layers with a thickness under five newtonmeters and titanium dioxide as an alternative gate nonconductor comprise, respectively, the last two chapters.			
14. SUBJECT TERMS German, UNIBW, MOSFET, Silicon dioxide layers, Boron diffusion, Boron penetration, Hot carrier stress, Bias temperature stress		15. NUMBER OF PAGES	
		16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UNLIMITED

UNIVERSITÄT DER BUNDESWEHR MÜNCHEN
Fakultät für Elektrotechnik
(Institut für Physik)

Gateisolatoren für MOS-Feldeffekttransistoren

Thomas Pompl

Vorsitzender des Promotionsausschusses: Prof. Dr. H. Baumgärtner

1. Berichterstatter: Prof. Dr. I. Eisele
2. Berichterstatter: Prof. Dr. K. Hoffmann

Tag der Prüfung 25.10.2000

Mit der Promotion erlangter akademischer Grad:
Doktor-Ingenieur
(Dr.-Ing.)

Neubiberg, den 30.10.2000

20021122 160

AQ F03-02-0397

**Reproduced From
Best Available Copy**

**Copies Furnished to DTIC
Reproduced From
Bound Originals**

Inhaltsverzeichnis

Zusammenfassung	v
1 Einleitung	1
2 Charakterisierungsverfahren	5
2.1 Ladungen und Haftstellen im Oxid – „Oxide Traps“	5
2.1.1 Die CV-Messung	7
2.1.2 „Charge Pumping“	11
2.2 Charakterisierung am MOS-Feldeffekttransistor	15
2.2.1 Einfluß der Gateoxidqualität auf wichtige Transistoreigenschaften	15
2.2.2 „Hot Carrier Stress“	18
2.2.3 „Bias Temperature Stress“	21
2.3 Zuverlässigkeitsuntersuchungen	22
2.3.1 Elektrische Streßmethoden	22
2.3.2 Beschleunigungsmodelle	23
2.3.3 Statistische Datenauswertung	26
2.3.4 Extrapolation der Lebensdauer auf Betriebsbedingungen	28
3 Nitridierte Gateoxide – Stickstoffbarriere gegen Bordiffusion	30
3.1 „Dual Workfunction CMOS“-Logik	30
3.2 Bordiffusion	31
3.3 Nitridierung von Siliziumdioxid	33
3.3.1 Thermische Nitridierung	33
3.3.2 Implantation von Stickstoff	33
3.3.3 Vergleich zwischen NO-Nitridierung und Stickstoffimplantation in das Siliziumsubstrat	35

3.3.4	Zusammenhang zwischen Bordiffusion und „Negative Bias Temperature Stress“	41
3.4	Zusammenfassung	44
4	Das Tunneln von Ladungsträgern durch eine dünne Siliziumdioxid-schicht	45
4.1	Tunnelstromdichte	45
4.1.1	Direktes Tunneln und Fowler-Nordheim-Tunneln	45
4.1.2	Berechnung mit einem Transmissionsmatrixansatz.....	46
4.2	Gateleckströme in „Dual Workfunction CMOS“-Technologien	47
4.2.1	Substrat in Akkumulation	47
4.2.2	Substrat in Inversion	51
4.3	Das Tunneln von Elektronen am Beispiel des PMOS und negative Gatespannungen	53
4.3.1	„Interface State Injection Model“	53
4.3.2	Berechnung mit dem Transmissionsmatrixansatz	55
4.3.3	Berechnung mit der Fowler-Nordheim-Gleichung	57
4.4	Weitere Gateleckstrommechanismen am Beispiel des PMOS und negative Gatespannungen	58
4.4.1	„Trap-Assisted Tunneling“	58
4.4.2	„Stress-Induced Leakage Current“	61
4.4.3	Leckstrom nach einem „Soft Breakdown“	63
4.4.4	Auswirkungen auf den PMOSFET	64
4.5	Gateleckstrommechanismen für den NMOS und positive Gatespannungen	64
4.6	Gateleckstrom aufgrund der Verformung der Potentialbarriere durch positive Oxidladungen	67
4.6.1	Potential einer Punktladung im Oxid	67
4.6.2	Positive Ladungen im Oxid	69
4.7	Auswirkung auf die Zuverlässigkeitsanalyse von Oxiden	70
4.7.1	E-Modell	71
4.7.2	I/E-Modell	72
4.7.3	Zusammenfassung der Diskussionsergebnisse	76
4.8	Zusammenfassung	77

5 Zuverlässigkeitsanalysen an Siliziumdioxidschichten mit Dicken unter 5 nm	79
5.1 Elektrische Degradation	79
5.1.1 „Stress-Induced Leakage Current“	79
5.1.2 „Soft Breakdown“	82
5.1.3 Der dielektrische Durchbruch – „Hard Breakdown“	84
5.1.4 Elektrischer Streß mit konstanter Spannung	86
5.2 Einfluß von „Soft Breakdown“ auf MOS-Feldeffekttransistoren	88
5.2.1 Gateleckstrom und Signalrauschen	88
5.2.2 Streßexperimente an MOS-Feldeffekttransistoren	88
5.2.3 „Gate-Induced Drain Leakage“ bei NMOSFETs	89
5.2.4 Modell für den Einfluß des „Soft Breakdown“ auf „Gate-Induced Drain Leakage“	91
5.2.5 „Soft Breakdown“ in unterschiedlichen Bereichen des Feldeffekttransistors	94
5.2.6 Einfluß auf die Funktionalität eines NMOSFETs	97
5.2.7 Einfluß auf die Funktionalität eines PMOSFETs	97
5.3 Modell des selbstlimitierenden „Soft Breakdown“-Pfades	97
5.4 Beschleunigung der Zeit bis zum „Soft Breakdown“ im Vergleich mit dem dielektrischen Durchbruch	99
5.4.1 Beschleunigung in Abhängigkeit vom elektrischen Feld	100
5.4.2 Beschleunigung in Abhängigkeit von der Temperatur	102
5.4.3 Analyse des „Soft Breakdown“	105
5.4.4 Auswirkungen auf die Zuverlässigkeitscharakterisierung	109
5.5 Zusammenfassung und Ausblick auf die zukünftigen Prozeßgenerationen	112
6 Titandioxid als alternatives Gatedielektrikum	114
6.1 Eigenschaften und Abscheidung von Titandioxid	114
6.1.1 Materialeigenschaften	114
6.1.2 Abscheideverfahren	114
6.2 Prozeßkonzepte mit Titandioxid als Gatedielektrikum	116
6.2.1 Basisprozeß	116
6.2.2 Prozeß mit Nitridschichten	117

6.2.3	Prozeß mit einem Metallgate	117
6.3	Prozeßtests und Schichtanalysen	118
6.4	Implementierung in einen 0,25µm-CMOS-Prozeß	120
6.4.1	Physikalische Analyse der Schichtdicken und der Gatekante	120
6.4.2	Eingangskennlinien von Feldeffekttransistoren	122
6.4.3	Leckströme des Dielektrikums	124
6.4.4	Strukturelle Schichteigenschaften des Dielektrikums	127
6.4.5	Charakterisierung mittels Photoemissionsmikroskopie	128
6.4.6	Abschätzung für die Dielektrizitätskonstante	129
6.5	Zusammenfassung	131
7	Anhang A: MOS-Strukturen und MOSFET	133
7.1	Das MOS-System – Energien und Potentiale	133
7.2	Eingangskennlinien eines MOSFETs – wichtige Größen und Begriffe	134
7.3	Ladungsträgerbeweglichkeiten und effektives elektrisches Feld	134
7.4	„Buried Channel p-FET“	135
7.5	„Gate-Induced Drain Leakage“	139
7.6	Kurzkanaleffekte	140
8	Anhang B: Statistik	141
8.1	Allgemeines zur Statistik	141
8.2	Die Weibull-Verteilung	141
8.3	Die Poisson-Verteilung	142
	Literaturverzeichnis	144
	Verzeichnis der Abkürzungen	154
	Verzeichnis der Symbole	156
	Danksagung	161

Zusammenfassung

Mit der thermischen Nitridierung in NO-Atmosphäre und einer Stickstoffimplantation in das Siliziumsubstrat wurden im ersten Teil der Arbeit zwei völlig unterschiedliche Nitridierungskonzepte miteinander verglichen. Im Gegensatz zu Darstellungen in der Fachliteratur wurde festgestellt, daß Stickstoff in der Grenzfläche zwischen SiO_2 und Substrat nicht automatisch zu einer Verbesserung der Hot-Carrier-Stabilität führt. Die bessere Stabilität der Grenzfläche, welche im allgemeinen als Grund für eine höhere Hot-Carrier-Stabilität angegeben wird, muß dafür gegenüber weiteren Degradationseinflüssen dominant werden. Im Fall des thermischen Oxides mit Stickstoffimplantation in das Substrat ergibt sich eine Stabilität gegenüber der Generation von Grenzflächenzuständen, die der eines reinen thermischen Oxides entspricht. Dies deutet auf eine gleiche Struktur der Grenzfläche hin und wird durch die Analyse der Beweglichkeit der Minoritätsladungsträger bei hohen elektrischen Feldern bestätigt. Die Stabilität des PMOS-Feldeffekttransistors gegenüber „Negative Bias Temperature Stress“ (NBTS) ist für unterschiedliche Nitridierungskonzepte verschieden. Dies kann erklärt werden, wenn berücksichtigt wird, daß Bor im Oxid die NBTS-Stabilität verschlechtert und die Borkonzentration durch die Struktur der Stickstoffbarriere beeinflusst wird.

Im zweiten Teil der Arbeit wurde gezeigt, daß unter Berücksichtigung der quantenmechanischen Transmission und Reflexion von Elektronen die Gateleckstromdichte durch eine dünne SiO_2 -Schicht bei hohen Spannungen sehr gut berechnet werden kann. Dabei sind für die verschiedenen Fälle, die bei Dual-Workfunction-CMOS unterschieden werden müssen, die korrekten Potentialverhältnisse zu betrachten. Für den PMOS und negative Gatespannungen wurde ein neues Modell eingeführt, das als „Interface State Injection Model“ bezeichnet wurde. Die daraus berechnete Tunnelstromdichte für Elektronen erklärt nicht nur den Gateleckstrom bei hohen Spannungen, sondern auch die Stromdichte am Diffusionskontakt einer diffusionsbegrenzten PMOS-Kapazität oder eines PMOS-Feldeffekttransistors. Aus der Analyse der gemessenen Stromdichten für diesen Fall, konnten Rückschlüsse auf die Mechanismen gewonnen werden, die „Trap-Assisted Tunneling“, „Stress-Induced Leakage Current“ und „Soft Breakdown“ bestimmen. Insbesondere wurde gezeigt, daß „Trap-Assisted Tunneling“ im wesentlichen auf einen inelastischen Transport von Elektronen im Oxid zurückzuführen ist. Zudem ergaben diese Untersuchungen, daß der Strom am Wannenkontakt des NMOS im Fall positiver Gatespannungen durch das Tunneln von Valenzbandelektronen erklärt werden kann. Die Potentialbarriere für das in der Literatur diskutierte „Valence Band Injection Model“ ist hierfür nur geringfügig zu korrigieren. Diese Korrektur wird durch das eingeführte „Interface State Injection Model“ begründet. Abschließend wurde im zweiten Teil der Arbeit gezeigt, daß die Verformung der

Potentialbarriere durch das Coulombpotential von positiven Ladungen im Oxid nicht als maßgebende Ursache zur Erklärung des Gateleckstromes aufgrund von „Trap-Assisted Tunneling“ herangezogen werden kann.

Der dritte Teil der Arbeit befaßt sich hauptsächlich mit dem Phänomen des „Soft Breakdown“. Es wurde erstmals gezeigt, daß der „Soft Breakdown“ zu einer unmittelbaren Degradation von NMOS-Feldeffekttransistoren führen kann. Diese tritt auf, wenn der Ort des „Soft Breakdown“ im Überlappbereich von Gate und Drain lokalisiert ist. Die Degradation äußert sich in einem stark erhöhten Off-Strom des Transistors, der durch einen erhöhten „Gate-Induced Drain Leakage“-Strom verursacht wird. Ein Modell zur Erklärung dieses Effektes wurde dargelegt. Dieses basiert auf der Anreicherung von negativen Ladungen im Bereich des „Soft Breakdown“. Das Modell begründet auch, warum eine vergleichbare Degradation im Fall eines PMOS-Feldeffekttransistors nicht gemessen wird. Ausgehend von diesem Modell wurde für den „Soft Breakdown“ das Modell eines selbstlimitierenden Strompfades für den Transport von Elektronen entwickelt. Dieses Modell erklärt die Fluktuationen im Gateleckstrom nach einem „Soft Breakdown“ und konnte durch experimentelle Ergebnisse untermauert werden. Die Beschleunigung der elektrischen Degradation dünner Oxide wurde in Abhängigkeit vom elektrischen Feld und von der Temperatur analysiert. Dabei erfolgte eine getrennte Charakterisierung für den „Soft Breakdown“ und für den dielektrischen Durchbruch. Es wurde erstmals gezeigt, daß sich das Beschleunigungsverhalten des „Soft Breakdown“ deutlich von dem des dielektrischen Durchbruches unterscheiden kann. Damit steht fest, daß beide Durchbruchereignisse stets getrennt voneinander charakterisiert werden müssen. Die experimentellen Daten für den „Soft Breakdown“ (SBD) zeigen eine Analogie zum thermochemischen E-Modell. Aufgrund dieser Analogie wurde eine Modellvorstellung für die Ausbildung eines SBD-Pfades entwickelt. Diese legt das Aufbrechen von H-Si-Bindungen und H-O-Bindungen unter dem Einfluß des elektrischen Feldes zugrunde. Die Fragmente bilden die Haftstellen für den Ladungsträgertransport. Eine Anreicherung von negativen Ladungen im Bereich des „Soft Breakdown“ bzw. im SBD-Pfad kann mit dieser Modellvorstellung erklärt werden, wenn die Emission von Elektronen durch Si \cdot - und O \cdot - eine geringere Reaktionsrate aufweist, als das Besetzen von H $^+$, Si \cdot und O \cdot mit einem Elektron.

Im vierten Teil der Arbeit wurde ein neues Verfahren zur Abscheidung von Titandioxid vorgestellt, das ausschließlich Standardprozesse verwendet. Das Verfahren ist geeignet, Titandioxid als Gatedielektrikum in verschiedenen Prozeßkonzepten für Logiktechnologien zu verwenden. Es wurde einer der ersten Versuche charakterisiert, Titandioxid als Gatedielektrikum in einen kompletten CMOS-Prozeß bis zur ersten Metallisierungsebene zu implementieren. Untersuchungen mittels Transmissionselektronenmikroskopie weisen nach, daß Titandioxid im verwendeten Gesamtprozeß große kristalline Körner bildet. Die Analyse der Leckstromdichte an verschiedenen Teststrukturen ergibt, daß durch strukturelle Defekte sehr hohe Leckströme verursacht werden können. Diese strukturellen Defekte äußern sich in weiten Bereichen zwischen Körnern aus Titandioxid, die mit Siliziumdioxid gefüllt sind. Eine Verteilung von Orten mit sehr hohen Leckströmen, die den strukturellen Defekten entsprechen, wurde mittels Photoemissionsmikroskopie nachgewiesen. Die Defekte treten bevorzugt bei großflächigen planaren Kapazitäten auf und werden durch die Möglichkeit zur

mechanischen Streßrelaxation beeinflusst. Es wurden die Eingangskennlinien von NMOS-Kurzkanaltransistoren und PMOS-Kurzkanaltransistoren gezeigt. Die Transistoren besitzen infolge der noch großen effektiven Gateoxiddicke hohe Einsatzspannungen. Für PMOS-Kurzkanaltransistoren tritt ein Kink-Effekt auf, der durch eine zusätzliche laterale Ausdiffusion von Source und Drain verursacht wird. An den Isolationskanten entstehen dadurch parasitäre Kurzkanaltransistoren mit einer kleineren Einsatzspannung und einer kleineren Weite. Aus den Analysen ist insgesamt zu schließen, daß ein Reoxidationsprozeß nach der Gatestrukturierung die Verwendung von Titandioxid als Gatedielektrikum negativ beeinflusst. Die physikalischen Analysen des Gatedielektrikums und die elektrische Charakterisierung von Kapazitäten und Transistoren zeigen, daß die vollständige Implementierung in einen CMOS-Prozeß möglich ist. Es werden jedoch wesentliche prozeßtechnische Änderungen erforderlich sein, um einen stabilen CMOS-Prozeß mit einem Gatedielektrikum aus Titandioxid oder einem anderen Metalloxid mit einer hohen Dielektrizitätskonstanten zu realisieren.

1 Einleitung

Die ersten integrierten Schaltkreise wurden 1959 mit Bipolartransistoren hergestellt. Der Bipolartransistor nutzt den von W. H. Brattain und J. Bardeen 1947 entdeckten [1.1, 1.2] und von W. Shockley 1948 mit der Theorie des p-n-Überganges theoretisch begründeten Transistoreffekt [1.3, 1.4]. Die Forschungsgruppe der Bell Laboratories war ursprünglich auf der Suche nach der experimentellen Realisierung des Feldeffektes für ein verstärkendes Bauelement auf der Basis eines Elementhalbleiters. Als Halbleitermaterial wurde Germanium verwendet. Nach dem Verständnis des Transistoreffektes wurde erkannt, daß Silizium aufgrund der größeren Bandlücke der geeignetere Halbleiter wäre. Im Gegensatz zum Germanium war jedoch die Herstellung eines sehr reinen Einkristalls aus Silizium zunächst problematisch. Transistoren, die den Feldeffekt nutzen, wurden erst realisiert, als weitere Erkenntnisse über Oberflächenzustände beim thermisch oxidierten Silizium gewonnen wurden. Zu diesem Zeitpunkt waren Substrate aus einkristallinem Silizium bereits ein Standard in der Halbleiterindustrie. Intel produzierte 1970 ein SRAM („Static Random-Access Memory“) auf der Basis von MOS-Feldeffekttransistoren („Metal Oxide Semiconductor“), dessen Schaltgeschwindigkeit vergleichbar mit der des ersten schnellen SRAMs von Fairchild Semiconductors auf der Basis von Bipolartransistoren war. Dies war der Beginn des Erfolges der MOS-Technologie. MOS-Transistoren sind technologisch einfacher zu fertigen und die lateralen Dimensionen können wesentlich stärker verkleinert werden als im Fall von Bipolartransistoren.

Es stellte sich dabei nachträglich heraus, daß die ursprüngliche Wahl des Halbleiters Silizium für Bipolartransistoren entscheidend für den Erfolg der MOS-Technologie war. Eine der fundamentalsten Vorteile von Silizium ist die Ausbildung eines natürlichen Oxides, Siliziumdioxid, welches ausgezeichnete Eigenschaften besitzt. Dies sind eine hohe dielektrische Durchbruchfeldstärke, die thermische Stabilität bei hohen Temperaturen und eine sehr gute Grenzfläche zwischen Siliziumdioxid und dem Siliziumsubstrat, welche bei der thermischen Oxidation von einkristallinem Silizium entsteht. Letzteres ist eine sehr wichtige Voraussetzung für die MOS-Technologie, welche eine Schicht aus Siliziumdioxid zur Isolation des Gates in MOS-Feldeffekttransistoren verwendet. Der Gateisolator wird daher generell als „Gateoxid“ bezeichnet. Eine geringe Dichte an Grenzflächenladungen ermöglicht, daß der Feldeffekt nicht bereits an der Substratoberfläche elektrisch abgeschirmt wird. Diese wichtige Voraussetzung für die Nutzung des Feldeffektes wurde bereits 1947 von J. Bardeen erkannt. Die Eigenschaften von Siliziumdioxid und die gute Kontrollierbarkeit des thermischen Oxidationsprozesses führten dazu, daß das Gateoxid bislang als invariantes Element in der stetigen Weiterentwicklung der MOS-Technologie galt. Die Veränderung des Gateoxides bei der Entwicklung von Schaltungen mit immer kleineren Strukturdimensionen

beschränkte sich im wesentlichen auf eine Verkleinerung der Schichtdicke. In der Abbildung 1.1 ist die Verkleinerung der Strukturdimensionen und der Gateoxiddicke über die Zeit gezeigt. In der Abbildung 1.2 ist anhand des Integrationsgrades die Entwicklung der MOS-Technologie bis zur Gegenwart dargestellt. MOS-Feldeffekttransistoren erlangten nach ihrer Realisierung eine größere Bedeutung als die früher entwickelten Bipolartransistoren.

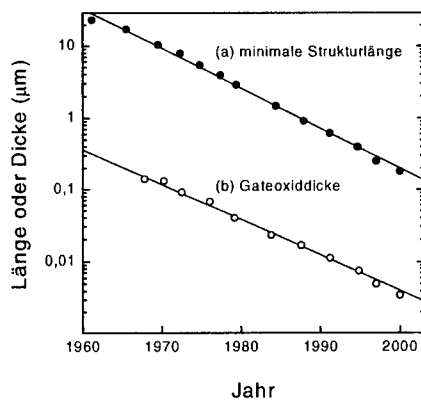


Abb. 1.1: Minimale Strukturlänge (a) und Gateoxiddicke (b) bis zur gegenwärtigen 0,18µm-Generation. Die Daten sind größtenteils aus [1.5] entnommen.

In der 0,18µm-Generation wurde für Logikprozesse erstmals eine technologische Veränderung des Gateisolators eingeführt. Die Kombination aus aktueller CMOS-Technologie (CMOS: Complementary MOS) mit Siliziumdioxiddicken deutlich unter 5 nm erforderte eine Nitridierung von Siliziumdioxid. Die Gründe hierfür werden zu Beginn des Kapitels 3 umfassend dargelegt. Dieses Kapitel befaßt sich mit thermisch nitridierten Oxiden, der Nitridierung durch Stickstoffimplantation in das Siliziumsubstrat vor der thermischen Oxidation und der Stickstoffimplantation in das Polysiliziumgate. Der Unterschied zu zahlreichen veröffentlichten Arbeiten über nitridierte Gateoxide ist ein direkter Vergleich zwischen dem neuesten thermischen Nitridierungsverfahren mit NO und der Nitridierung durch Stickstoffimplantation in das Siliziumsubstrat. Dabei werden Eigenschaften wie z.B. Hot-Carrier-Stabilität oder Bias-Temperature-Stabilität behandelt. Letztere wird auch im Vergleich mit der Stickstoffimplantation in das Polysiliziumgate und mit einem in N₂O-Atmosphäre thermisch nitridierten Oxid untersucht.

Neben den technologischen Aspekten gewinnen für die nächsten Prozeßgenerationen physikalische Effekte an Bedeutung, die mit der fortschreitenden Abnahme der Oxiddicke verbunden sind. Für Oxiddicken kleiner ca. 6 nm wird z.B. der Transport von Elektronen im

Leitungsband von Siliziumdioxid ballistisch. Im Kapitel 4 wird die Tunnelstromdichte an Elektronen mittels quantenmechanischer Transmission und Reflexion beschrieben. Einen wichtigen Aspekt stellen dabei die Potentialverhältnisse dar, welche sich für NMOS- und PMOS-Strukturen bei Dual-Workfunction-CMOS-Technologien ergeben. Es wird erstmals ein Modell vorgestellt, das die Berechnung der Tunnelstromdichte für PMOS-Strukturen mit positiv dotiertem Gate und negativen Gatespannungen ermöglicht. Außerdem wird auf weitere Leckstrommechanismen eingegangen, die bei dünnen Gateoxiden bedeutsam werden.

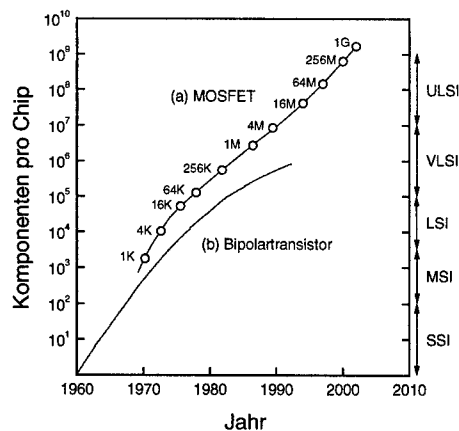


Abb. 1.2: Entwicklung des Integrationsgrades pro Chip für (a) MOSFET und (b) Bipolartransistor (Daten weitgehend aus [1.5]). Die Abkürzungen der Integrationsgrade sind: „small-scale integration“ (SSI), „medium-scale integration“ (MSI), „large-scale integration“ (LSI), „very-large-scale integration“ (VLSI) und „ultralarge-scale integration“ (ULSI).

Die Zuverlässigkeit von Oxiden wird mit abnehmender Oxiddicke durch neue Formen der Degradation beeinflusst. Die Untersuchung der Oxidzuverlässigkeit beschränkte sich bislang im wesentlichen auf die Bestimmung der Lebensdauer bis zu einem dielektrischen Durchbruch. Ab einer Oxiddicke von ca. 5 nm tritt eine neue Form der Degradation auf, für die sich in der Literatur die Bezeichnungen „Soft Breakdown“ und „Quasi-breakdown“ etabliert haben. Im Kapitel 5 werden neue Erkenntnisse bezüglich Eigenschaften und Auswirkungen des „Soft Breakdown“ dargelegt. Diese beantworten unter anderem die bislang offene Frage, ob der „Soft Breakdown“ Auswirkungen auf die Zuverlässigkeit von integrierten Schaltungen mit MOS-Technologie hat. Zudem wird eine genauere Vorstellung vom Phänomen „Soft Breakdown“ gewonnen, als sie bisher in der Literatur gegeben war. Die Charakterisierung der Lebensdauer von Gateoxiden unter Berücksichtigung des „Soft Breakdown“ wird ausführlich analysiert.

Die weitere Verkleinerung der Schichtdicke des Gateoxides stößt bereits in den nächsten 10 Jahren an technologische und physikalische Grenzen. In Tabelle 1.1 ist diese Entwicklung gezeigt. Neben der Lithographie zur Strukturierung von planaren Schichten wird der Gateisolator immer häufiger als weiteres sehr kritisches Entwicklungselement der MOS-Technologie genannt. Die exakte Festlegung der realisierbaren minimalen Schichtdicke für Siliziumdioxid wird noch kontrovers diskutiert und ist auch abhängig von Aspekten der Zuverlässigkeit. Aus physikalischer Sicht können jedoch zwei atomare Monolagen keine Schicht mit der Stöchiometrie SiO_2 und entsprechender Dielektrizitätskonstanten darstellen. Bereits ab einer Dicke von ca. 2 nm verliert die Schicht ihre isolierende Eigenschaft aufgrund des direkten Tunnelns von Ladungsträgern. Die Suche nach alternativen Dielektrika mit höheren Dielektrizitätskonstanten, welche Siliziumdioxid als Gateisolator ersetzen können, wird daher ein immer wichtigerer Faktor in der Technologieentwicklung. Das Kapitel 6 stellt einen der weltweit ersten Versuche vor, Titandioxid als Gatedielektrikum in einen kompletten Dual-Workfunction-CMOS-Prozess zu integrieren.

Zu Beginn der Arbeit werden im Kapitel 2 Charakterisierungsverfahren für Dielektrika, insbesondere Siliziumdioxid, beschrieben, die größtenteils auch für die Charakterisierung der nitridierten Gateoxide im Kapitel 3 verwendet wurden. Dabei wird u.a. ein Verfahren zur Bestimmung der energieabhängigen Grenzflächenzustandsdichte mittels „Charge Pumping“ verbessert. Das Kapitel 2.3 enthält Informationen über die Zuverlässigkeitsuntersuchung von Dielektrika. Diese werden in den Kapiteln 4 und 5 für die Diskussion und das Verständnis der Ergebnisse benötigt. Im Kapitel 4 betrifft dies vor allem die Diskussion der Auswirkungen der gewonnen Ergebnisse über die Tunnel- und Leckströme auf die Zuverlässigkeitsanalyse von Oxiden. Im Kapitel 5 werden die Kenntnisse generell für die Charakterisierung des „Soft Breakdown“ vorausgesetzt.

Jahr der ersten Produktion	1997	1999	2002	2005	2008	2011
Technologiegeneration minimale Strukturbreite (nm)	250	180	130	100	70	50
äquivalente Gateoxiddicke T_{ox} (nm)	4-5	3-4	2-3	1.5-2	<1.5	<1.0

Tab. 1.1: Skalierung der Gateoxiddicke für die zukünftigen Technologiegenerationen [1.6]. Es ist die äquivalente Gateoxiddicke angegeben, d.h. für die maximale Dicke T eines beliebigen Dielektrikums mit der relativen Dielektrizitätskonstanten ϵ gilt: $T = T_{\text{ox}} \cdot \epsilon / \epsilon_{\text{SiO}_2}$, mit $\epsilon_{\text{SiO}_2} = 3,9$. Ab der 70nm-Generation kann noch keine technologische Lösung für den Gateisolator angegeben werden.

2 Charakterisierungsverfahren

In diesem Kapitel werden grundlegende Verfahren zur elektrischen Charakterisierung von Dielektrika, insbesondere Gateoxiden behandelt. Dabei liegt der Schwerpunkt nicht auf einer vollständigen Beschreibung der Meßverfahren und deren Theorie. Es werden zum einen Aspekte behandelt, die zum Verständnis von Ergebnissen beitragen, und zum anderen sollen die zu untersuchenden Systeme dargelegt werden.

2.1 Ladungen und Haftstellen im Oxid – „Oxide Traps“

Fremdatome und Fehler in der Bindungsstruktur von Oxiden verursachen Defekte, welche die periodische Anordnung des Bindungsnetzwerkes lokal stören. Liegt das Energieniveau eines Defektes in der Bandlücke des Oxides, so kann dieser Defekt entweder als ein Rekombinationszentrum oder als eine Haftstelle für freie Ladungsträger wirken, die in das Oxid injiziert wurden. Aufgrund des Haftens von Ladungsträgern spricht man in der Literatur von „traps“ (trap: Falle). Da Siliziumdioxid seit Jahrzehnten als Gatedielektrikum in MOS-Feldeffekttransistoren verwendet wird, sind in der Praxis die Analysen und Meßmethoden insbesondere für das MOS-System ausgearbeitet. Daher beziehen sich alle folgenden Abschnitte auf Siliziumdioxid und MOS-Strukturen.

Die Haftstellen können auf drei verschiedene Arten klassifiziert werden [2.1]. Das Konzept der Ionisationsenergie unterscheidet die Haftstellen je nach Potentialdifferenz zwischen dem Energieniveau der Haftstelle und der jeweiligen Bandkante. Im Fall von Elektronen ist dies die Leitungsbandkante und für Löcher entsprechend die Valenzbandkante des Oxides. Trotz der amorphen Struktur wird für Siliziumdioxid die Terminologie des Bändermodells verwendet. Die Bandlücke beträgt bei 300 K ca. 9 eV [2.2]. Es wird zwischen „seichten“ und tiefen Haftstellen unterschieden. Seichte Haftstellen besitzen eine sehr niedrige Ionisationsenergie zwischen 0,010 eV und 0,10 eV, was in etwa der Energie eines Phonons in SiO₂ entspricht. Die Ionisationsenergie tiefer Haftstellen beträgt ein Vielfaches dessen [2.1].

Die zweite Klassifizierungsmethode berücksichtigt den Ladungszustand der Haftstelle. Donatorartige Haftstellen sind z.B. neutral, wenn sie mit einem Elektron besetzt sind, und positiv geladen, wenn nicht. Akzeptorartige Haftstellen sind in bezug auf Elektronen negativ geladen, wenn sie besetzt sind, und neutral, falls dies nicht der Fall ist. Diese Klassifizierung ist schwierig, wenn derselbe Defekt auch für Löcher relevant ist. Eine Übersicht ist in Tabelle 2.1 gegeben.

Die dritte Art der Klassifizierung führt die Definition eines Wirkungsquerschnittes für Haftstellen σ_T ein. Betrachtet man die Verformung der Bandkanten, welche durch die lokale Veränderung des elektrostatischen Feldes verursacht wird, dann bezieht sich der Wirkungsquerschnitt auf die Öffnung dieses Potentialtopfes. Je nach der Form des Potentialtopfrandes wird unterschieden zwischen [2.1]

- einer Haftstelle mit anziehendem Coulombpotential, wenn $10^{-14} < \sigma_T < 10^{-12} \text{ cm}^2$,
- einer Haftstelle mit abstoßendem Coulombpotential, wenn $10^{-22} < \sigma_T < 10^{-18} \text{ cm}^2$, und
- einer neutralen Haftstelle, wenn $10^{-18} < \sigma_T < 10^{-14} \text{ cm}^2$ ist.

Beteiligter Ladungsträger	Symbol der leeren Haftstelle	Coulombpot. der Haftstelle	Besetzen der Haftstelle	σ_T variiert mit	Freisetzen des Ladungsträgers	Typ der Haftstelle
e^-	X^+	anziehend	$X^+ + e^- \rightarrow X^0$	$E_{ox}^{-1/2}, T^{-a}$	$X^0 \rightarrow X^+ + e^-$	Donator
e^-	X^0	neutral	$X^0 + e^- \rightarrow X^-$	E_{ox} (schwach)	$X^- \rightarrow X^0 + e^-$	Akzeptor
e^-	X^-	abstoßend	$X^- + e^- \rightarrow X^{2-}$	E_{ox}, T	$X^{2-} \rightarrow X^- + e^-$	
h^+	X^+	abstoßend	$X^+ + h^+ \rightarrow X^{2+}$	T^{-b}	$X^{2+} \rightarrow X^+ + h^+$	
h^+	X^0	neutral	$X^0 + h^+ \rightarrow X^+$	-	$X^+ \rightarrow X^0 + h^+$	Donator
h^+	X^-	anziehend	$X^- + h^+ \rightarrow X^0$	T^{-b}	$X^0 \rightarrow X^- + h^+$	Akzeptor

Tab. 2.1: Klassifizierung von Haftstellen nach Ladungstyp und Wirkungsquerschnitt gemäß [2.1] und [2.3].

Die Charakterisierung des Gateoxides beschränkt sich im allgemeinen auf die festen Oxidladungen und die Grenzflächenzustände. Aufgrund des Strukturwechsels beinhaltet die Grenzfläche zwischen SiO_2 und einkristallinem Silizium Bindungsdefekte. In dieser Übergangsschicht, die in der Literatur als „structural transition layer“ bezeichnet wird, sind feste Oxidladungen lokalisiert, die nicht umgeladen werden können. Ionisiertes Silizium, das noch nicht oxidiert ist, wird als Ursache für diese festen Oxidladungen vermutet [2.4]. Als Grenzflächenzustände („interface traps“) werden Haftstellen bezeichnet, die in dieser Übergangsschicht lokalisiert sind. Die Quelle dieser Haftstellen sind nicht besetzte Bindungen („dangling bonds“) an der Siliziumoberfläche [2.4]. Diese Grenzflächenzustände beeinflussen die Eigenschaften eines MOS-Feldeffekttransistors und können durch den Reinigungsprozeß des Siliziumsubstrates vor der thermischen Oxidation des Gateoxides minimiert werden. Daher ist der Charakterisierung der Grenzflächenzustände große Aufmerksamkeit gewidmet worden. Zwei Meßmethoden, die hauptsächlich nur für die Bestimmung der Dichte der Grenzflächenzustände Anwendung finden, werden im folgenden behandelt. In beiden Fällen beschränkt sich die Meßmethode auf die Grenzflächenzustände mit Energieniveaus in der

Bandlücke des Siliziumsubstrates. Diese Zustände unterliegen somit einem Emissions- und Besetzungsvorgang durch Elektronen und Löcher. Bei einer Potentialbarriere Φ_b von ca. 3,1 eV zwischen den Leitungsbandkanten von SiO_2 und Silizium haben diese Niveaus also eine Ionisationsenergie von ca. 3,1 eV bis 4,2 eV bezüglich der Oxidleitungsbandkante. Aus der Dichte der Grenzflächenladungen Q_{it} berechnet sich die absolute Dichte der Grenzflächenzustände N_{it} zu:

$$N_{it} = \frac{Q_{it}}{q} \quad (2.1)$$

q ist die Elementarladung. Um die Energieverteilung der Grenzflächenzustände in der Bandlücke des Siliziums mit einzubeziehen, wird meist die energieabhängige Definition der Grenzflächenzustandsdichte verwendet [2.2]:

$$D_{it} = \frac{1}{q} \cdot \frac{dQ_{it}}{dE} \quad (2.2)$$

2.1.1 Die CV-Messung

Eine CV-Messung bestimmt die Kapazität einer MOS-Struktur in Abhängigkeit von der angelegten Gatespannung. Es wird zwischen drei Meßmethoden unterschieden, der Hochfrequenzmessung („high frequency“, HF-CV), der Niederfrequenzmessung („low frequency“, LF-CV) und der statischen CV-Messung.

Die statische CV-Messung

An einer MOS-Kapazität wird eine Spannungsrampe durchfahren. Die Spannungsrampe ändert den Zustand des Siliziumsubstrates an der Oberfläche zum Oxid von Akkumulation zu Inversion oder umgekehrt. Nach jedem Spannungsschritt wird das elektrodynamische Gleichgewicht abgewartet. Dies ist per Definition beim Unterschreiten eines bestimmten Verschiebestromes erreicht. Die Ladungsmessung an einer Elektrode ergibt dann die Kapazität des Systems in Abhängigkeit von der angelegten Spannung.

$$C(V) = \frac{Q(V)}{V} \quad (2.3)$$

Die Methode wird auch als Q-V-Methode bezeichnet [2.3].

LF-CV- und HF-CV-Messungen

In diesen Fällen wird an einer MOS-Kapazität eine Spannungsrampe durchfahren, die von einem frequenzabhängigen Spannungssignal mit kleiner Amplitude überlagert ist. Die Amplitude liegt dabei in etwa zwischen 40 mV und 100 mV. Das Kleinsignal mit der Frequenz f bewirkt, daß die Bandverbiegung frequenzabhängig moduliert wird. Damit moduliert auch die energetische Lage der Grenzflächenzustände bezüglich des Fermi-niveaus des Substrates. Es werden daher alle Grenzflächenzustände, deren Energieniveaus sich gerade

in der Nähe des Fermi-niveaus des Substrates befinden, mit der Periode $T = 1/f$ umbesetzt. Durch die beteiligten Kapazitäten entsteht ein imaginärer Anteil zur Impedanz des Schaltkreises gemäß

$$Z = R + \frac{1}{i\omega C} \quad (2.4)$$

Mit einem Impedanzmeter kann auf diese Weise die Kapazität des Systems in den unterschiedlichen Zuständen Akkumulation, Verarmung und Inversion vermessen werden. Ersatzschaltbilder für den Schaltkreis finden sich z.B. in [2.3].

Bei HF-CV-Messungen und LF-CV-Messungen sind folgende Unterschiede zu beachten. In starker Inversion ist die Bereitstellung von Ladungsträgern an der Grenzfläche abhängig von der Generation und Rekombination von Minoritätsladungsträgern. Die Zeitkonstante in starker Inversion liegt bei 0,01 s bis 1 s [2.3]. Damit wird nur für sehr kleine Frequenzen von ca. 10 Hz ein vollständiger Anstieg der Kapazität in Inversion erreicht. Wenn nur einfache Kapazitätsstrukturen oder isolationsbegrenzte Kapazitäten zur Verfügung stehen, kann dieses Problem eventuell mit der Verwendung von künstlichem Licht gelöst werden. Durch Photonen werden dann zusätzlich Minoritätsladungsträger generiert. Eine andere Möglichkeit bietet die CV-Messung an einer diffusionsbegrenzten Kapazität. Über das kontaktierte Diffusionsgebiet können ausreichend Minoritätsladungsträger zur Verfügung gestellt werden, so daß sogar bei sehr hohen Frequenzen die Oxidkapazität in Inversion erreicht wird.

HF-CV-Messungen werden üblicherweise bei 10 kHz bis 1 MHz durchgeführt. Die Zeitkonstante der Majoritätsladungsträger ist in Akkumulation und im Übergang zur Verarmung ca. 10^{-12} s [2.3]. Sie können also diesen Frequenzen folgen. Im Gegensatz dazu ist die Zeitkonstante für den Emissions- und Besetzungsvorgang von Grenzflächenzuständen größer, so daß deren Beitrag zur Kapazität bei diesen Frequenzen verschwindet. Das wird genutzt, um mit einer kombinierten Messung mit hohen und niedrigen Frequenzen die Grenzflächenzustandsdichte zu bestimmen. Für die Kapazität bei einer LF-CV-Messung $C_{1,f}$ gilt [2.3]:

$$\frac{1}{C_{1,f}} = \frac{1}{C_{ox}} + \frac{1}{C_s + C_{it}} \quad (2.5)$$

C_{ox} ist die Oxidkapazität. C_{it} ist die Kapazität der Grenzflächenzustände:

$$C_{it} \approx q \cdot D_{it} \quad (2.6)$$

C_s ist die Kapazität, welche durch das Halbleitersubstrat verursacht wird. Es gilt allgemein:

$$C_s = C_D + C_I \quad (2.7)$$

Dabei ist C_D die Kapazität der Verarmungszone und C_I ist die Kapazität der Inversionsladungen.

Im Fall der HF-CV-Messung gilt [2.3]:

$$\frac{1}{C_{HF}} = \frac{1}{C_{ox}} + \frac{1}{C_s} \quad (2.8)$$

Daraus ermitteln sich C_{it} und D_{it} wie folgt:

$$C_{it} = \left[\frac{1}{C_{LF}} - \frac{1}{C_{ox}} \right]^{-1} - \left[\frac{1}{C_{HF}} - \frac{1}{C_{ox}} \right]^{-1} \quad (2.9)$$

$$D_{it} = \frac{\Delta C}{q} \cdot \left[1 - \frac{C_{HF} + \Delta C}{C_{ox}} \right]^{-1} \cdot \left[1 - \frac{C_{HF}}{C_{ox}} \right]^{-1} \quad (2.10)$$

Dabei ist $\Delta C = (C_{LF} - C_{HF})$.

Wird das MOS-System von Akkumulation zu schnell in den Zustand der Inversion gebracht, dann können die Minoritätsladungsträger selbst dem frequenzunabhängigen Spannungssignal nicht folgen. Es stellt sich in der Verarmungszone kein elektrisches Gleichgewicht ein. Die Folge ist eine starke Verarmung („deep depletion“). Die Schrittrate der Spannungsrampe sollte in etwa 10-50 mV/s betragen. Ein Wert von 0,1 V/s ist bereits zu hoch [2.3]. Die Verläufe von HF-CV-Kurve, LF-CV-Kurve und für starke Verarmung sind jeweils in der Abbildung 2.1 schematisch dargestellt. Eine Erläuterung der Potentiale ist im Kapitel 7.1 des Anhanges A angegeben.

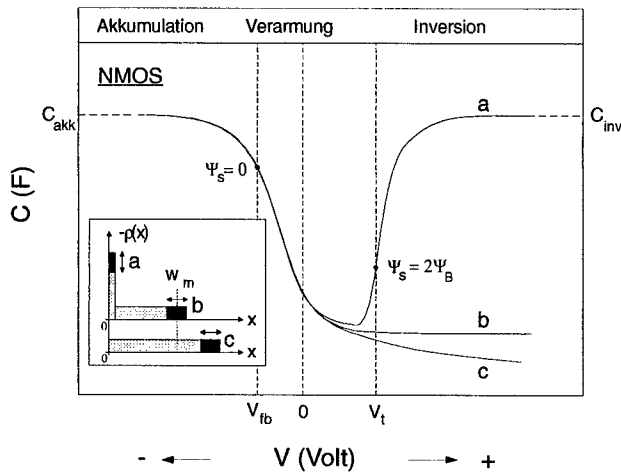


Abb. 2.1: Schematischer Verlauf einer LF-CV-Kurve (Fall a), einer HF-CV-Kurve (Fall b) und für starke Verarmung („deep depletion“) im Fall c. Im Abbildungsausschnitt sind die beteiligten Ladungen im invertierten Silizium für die Fälle a, b und c skizziert. p ist die Ladungsdichte und w_m ist die maximale Weite der Verarmungszone im elektrischen Gleichgewicht.

Zusammenhang zwischen der Gatespannung und der Bandverbiegung

Ein Zusammenhang zwischen der Gatespannung V_g und der Bandverbiegung an der Oberfläche des Siliziumsubstrates Ψ_s ermöglicht, daß der Grenzflächenzustandsdichte $D_{it}(V_g)$ die jeweilige Lage in der Bandlücke des Siliziums zugeordnet werden kann. Es ist [2.3]:

$$\frac{d\Psi_s}{dV_g} = 1 - \frac{C_{LF}}{C_{ox}} \quad (2.11)$$

$$\Psi_s(V_g) = \int_{V_0}^{V_g} \left[1 - \frac{C_{LF}}{C_{ox}} \right] \cdot dV + \Psi_s(V_0) \quad (2.12)$$

Da für die Flachbandspannung $\Psi_s(V_{fb}) = 0$ gilt, kann durch eine numerische Integration von Akkumulation bis zu V_{fb} und von V_{fb} bis zur Inversion der gesamte Verlauf von $\Psi_s(V_g)$ experimentell ermittelt werden. Hierfür muß aber die Flachbandspannung bereits bestimmt sein. Wenn dies nicht der Fall ist, kann für die Spannung V_0 auch ein beliebiger Wert in Akkumulation oder Inversion gewählt werden, wo die Bandverbiegung nur eine schwache Funktion von der Gatespannung ist. Somit ist dann der Fehler durch die Vernachlässigung der unbekannten Größe $\Psi_s(V_0)$ klein [2.3].

Dehnung der CV-Kurve durch Grenzflächenzustände

Wenn sich das System im Gleichgewicht befindet, egalisiert die Oberflächenladung die Ladung am Gate vollständig. Liegen Grenzflächenzustände vor, die eine zusätzliche Ladung Q_{it} beitragen, dann muß das System weniger Oberflächenladung zur Verfügung stellen. Damit wird auch die Bandverbiegung geringer. Um das System von Akkumulation nach Inversion zu bringen, ist daher ein größerer Spannungsbereich erforderlich. Dies führt zu einer Dehnung („stretch-out“) der CV-Kurve durch Grenzflächenzustände. Diese Dehnung wirkt sich auch auf eine HF-CV-Messung aus und ist von der in Gleichung 2.8 gemachten Aussage zu unterscheiden.

MOS-Kapazitäten mit Gates aus Polysilizium

In diesem Fall ist die Bandstruktur im Polysilizium zu berücksichtigen. Um einen Spannungsabfall zu vermeiden, werden Polysiliziumgates allerdings hoch dotiert. Unter Betriebsbedingungen eines MOSFETs sollte daher nur die Substratseite invertiert sein. Eine Überprüfung kann mit Hilfe des Verlaufes der LF-CV-Kurve in Inversion erfolgen. Liegt eine Verarmung im Polysilizium vor, dann erreicht die Kapazität in Inversion zunächst nicht den Maximalwert der Oxidkapazität. Erst bei einsetzender Inversion im Polysilizium steigt die Kapazität mit steigendem Gatespannungsbetrag wieder an.

CV-Messung an sehr dünnen Oxidschichten

Im Gegensatz zu einer Metallelektrode besitzt die Oberflächenladung eines Halbleiters eine endliche Ausdehnung. Der Abstand des Ladungsschwerpunktes von der Oberfläche läßt sich berechnen [2.5] und liegt im Bereich von 2 nm bis 10 nm [2.6, 2.7]. Die elektrische

Oxiddicke setzt sich damit aus der physikalischen Oxiddicke und dem Abstand des Ladungsschwerpunktes von der Grenzfläche zwischen SiO_2 und Silizium zusammen. Wird als Gateelektrode Polysilizium verwendet, dann ist dieser Effekt auch auf der Gateseite zu berücksichtigen. Der Abstand des Ladungsschwerpunktes ist abhängig von der Potentialdifferenz zwischen Kathode und Anode. Daher wird für Gateoxide mit Schichtdicken kleiner als ca. 5 nm in Akkumulation keine Sättigung der Kapazität erreicht, wie das für dickere Oxide der Fall ist. Eine einfache Bestimmung der Oxiddicke aus der Kapazität in Akkumulation ist nicht mehr möglich. Im Fall von sehr dünnen Oxiden, ca. 3 nm und weniger, beeinflusst der direkte Tunnelstrom die CV-Messung. Eine statische CV-Messung ist nicht mehr durchführbar. Eine frequenzabhängige CV-Messung ist nur noch im Bereich der Verarmung sinnvoll. Die Auswirkungen auf die Bestimmung der Gateoxiddicke wird seit über 10 Jahren untersucht [2.7, 2.8, 2.9, 2.10, 2.11, 2.12, 2.13]. Simulationsmodelle wurden entwickelt, welche sowohl die quantenmechanische Berechnung des Ladungsschwerpunktes als auch den Tunnelstrom berücksichtigen [2.11, 2.12, 2.13]. Mit diesen Modellen ist die Bestimmung der physikalischen Oxiddicke aus dem Vergleich der simulierten CV-Kurve mit dem gemessenen Anteil der CV-Kurve möglich.

2.1.2 „Charge Pumping“

Für diese Meßmethode werden Transistorstrukturen oder diffusionsbegrenzte Kapazitäten benötigt. Brugler und Jespers haben die Methode 1969 erstmals vorgestellt [2.14]. Der prinzipielle Meßaufbau ist in der Abbildung 2.2 gezeigt. Mit Hilfe eines Pulsers wird am Gatekontakt ein frequenzabhängiges Spannungssignal angelegt, wodurch das Substrat zwischen den Zuständen Inversion ($V_g > V_i$) und Akkumulation ($V_g < V_{fb}$) wechselt. In Inversion findet eine Besetzung der Grenzflächenzustände mit Minoritätsladungsträgern statt. Beim Übergang von Inversion zu Akkumulation werden die freien Minoritätsladungsträger unter dem Einfluß einer kleinen elektrischen Spannung wieder zurückgeführt, die am Diffusionsgebiet angelegt wird. Die Minoritätsladungsträger in den Grenzflächenzuständen rekombinieren mit Majoritätsladungsträgern. Somit werden permanent Ladungen eines Vorzeichens vom Diffusionsgebiet zur Wanne „geschaufelt“. Aus anderer Sicht werden Ladungen umgekehrten Vorzeichens von der Wanne in das Diffusionsgebiet gebracht. Die „Schaufel“ für die Ladungen sind die Grenzflächenzustände. In der Literatur hat sich für diesen Effekt die ursprüngliche Beschreibung „Charge Pumping“ etabliert. Diese Bezeichnung wird daher im weiteren beibehalten und gegebenenfalls mit CP abgekürzt. Der gemessene Strom an der Wanne ist umgekehrt äquivalent zum Strom, welcher am Diffusionskontakt gemessen wird. Im Fall von Transistoren werden Source und Drain auf das gleiche Potential gelegt. Der Charge-Pumping-Strom I_{CP} berechnet sich aus der Gatefläche A, der Frequenz f und der Dichte der Grenzflächenzustände N_H wie folgt:

$$I_{CP} = N_H \cdot A \cdot q \cdot f \quad (2.13)$$

Es gibt zwei grundlegende Meßvarianten. Bei der Variante mit variabler Pulsamplitude wird das Grundniveau des Spannungspulses V_{bottom} konstant gehalten, während das obere Niveau V_{top} solange erhöht wird, bis es die Flachbandspannung und die Einsatzspannung übertrifft. Der CP-Strom steigt stetig an, bis er eine Sättigung erreicht.

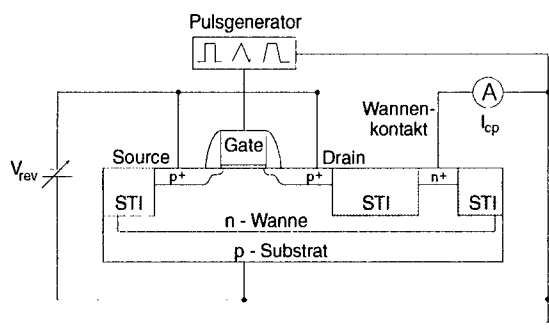


Abb. 2.2: Prinzipieller Meßaufbau für Charge-Pumping am Beispiel eines PMOSFETs. Die kleine Spannung am Diffusionsgebiet wird in der Literatur als „reverse voltage“ V_{rev} bezeichnet.

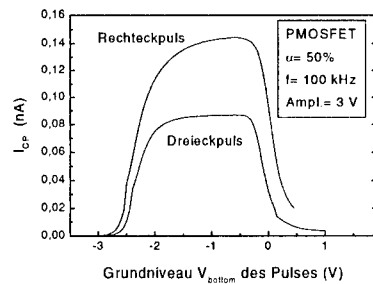


Abb. 2.3: Vergleich von Rechteckpuls und Dreieckpuls bei einer Pulsamplitude ΔV_g von 3 V. Die Charge-Pumping-Ströme wurden am gleichen PMOSFET gemessen. Die Flanken liegen bei $V_{th} - \Delta V_g$ und V_t . Für einen NMOSFET sind die Flanken bei $V_t - \Delta V_g$ und V_{th} lokalisiert.

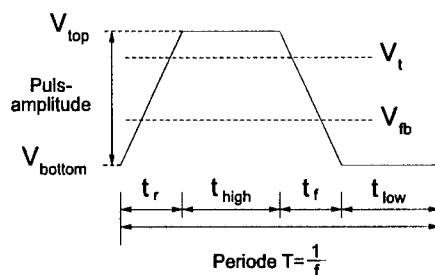


Abb. 2.4: Trapezförmiger Puls mit den Bezeichnungen der einzelnen Spannungsniveaus und Zeiten. Die Lage der Einsatzspannung V_t und der Flachbandspannung V_{th} ist für den Fall skizziert, der zur Erzeugung des vollen Charge-Pumping-Stromes führt.

Für die zweite Variante wird die Pulsamplitude konstant gehalten und das Grundniveau in einer Spannungsrampe durchfahren. Der CP-Strom nimmt, wie in der Abbildung 2.3 gezeigt, die Form einer Glockenkurve an. Im folgenden wird nur diese zweite Variante betrachtet.

Pulsform und Geometriekomponente

Ein Dreieckspuls weist den Nachteil auf, daß das System nur für hohe Pulsamplituden lange genug in Akkumulation oder Inversion verbleibt, um alle Grenzflächenzustände umzubesetzen [2.15]. Der Unterschied zwischen einem Rechteckpuls und einem Dreieckspuls ist in der Abbildung 2.3 gezeigt. Im Fall eines Rechteckpulses wird bei Geometrien mit einem Verhältnis von Weite zu Länge $W/L \leq 1$ ein außerordentliches Strommaximum im Plateau des CP-Stromes beobachtet. Dieses Strommaximum, das als Geometriekomponente bekannt wurde, tritt nicht auf für Geometrien mit $W/L > 1$ bzw. andere Pulsformen [2.15]. Unabhängig von diesem Effekt ist festzustellen, daß wegen der steilen Flanken eines Rechteckpulses das System bei dieser Pulsform sehr schnell zwischen Akkumulation und Inversion wechselt. Aufgrund dieser Nachteile von Dreieckspuls und Rechteckpuls wird ein trapezförmiger Puls, wie er in der Abbildung 2.4 skizziert ist, bevorzugt. Die CP-Methode ist nicht auf Kurzkanaltransistoren oder Strukturen mit kleiner Distanz zwischen den Diffusionsgebieten beschränkt [2.15].

Modifiziertes Verfahren zur Bestimmung der Grenzflächenzustandsdichte

Eine Methode zur Bestimmung der Grenzflächenzustandsdichte D_{it} mit dem Charge-Pumping-Meßverfahren wurde von Groeseneken vorgestellt [2.15]. Dieses Verfahren basiert auf dem in der gleichen Arbeit vorgestellten Modell für Charge-Pumping. Es nutzt die Frequenzabhängigkeit der Emissionszeiten von Elektronen und Löchern im Fall eines Dreieckspulses. Als Emissionszeit ist der Zeitabschnitt bezeichnet, innerhalb dem die Ladungsträger aus Grenzflächenzuständen in das Leitungsband bzw. Valenzband emittiert werden können. Aufgrund der im letzten Abschnitt besprochenen Nachteile eines Dreieckspulses wurde die Meßmethode modifiziert, um einen trapezförmigen Puls verwenden zu können. Mit Hilfe der Theorie für Charge-Pumping wird im folgenden gezeigt, daß hierfür der trapezförmige Puls bestimmte Eigenschaften erfüllen muß, so daß eine definierte Abhängigkeit der Emissionszeiten von der Frequenz erhalten bleibt.

Die Amplitude des Pulses ist:

$$\Delta V_g = V_{top} - V_{bottom} \quad (2.14)$$

Für die Emissionszeiten von Löchern und Elektronen gilt [2.15] (vgl. Abbildung 2.4):

$$t_{em,h} = \frac{|V_{fb} - V_t|}{|\Delta V_g|} \cdot t_r \quad t_{em,e} = \frac{|V_{fb} - V_t|}{|\Delta V_g|} \cdot t_r \quad (2.15)$$

Entscheidend ist nun, daß die Verhältnisse von t_r zu t_{high} und t_r zu t_{low} ebenso wie der „duty cycle“ α konstant gehalten werden.

$$\alpha = f \cdot (t_r + t_{high}) = \text{const.} \quad (2.16)$$

$$\frac{t_{\text{low}}}{t_f} = \text{const.} = k_i \quad \rightarrow \quad t_i = \frac{1 - \alpha}{f \cdot (1 + k_i)} \quad (2.17)$$

$$\frac{t_{\text{high}}}{t_r} = \text{const.} = k_r \quad \rightarrow \quad t_r = \frac{\alpha}{f \cdot (1 + k_r)} \quad (2.18)$$

Für den Charge-Pumping-Strom gilt [2.15]:

$$I_{\text{CP}} = \bar{D}_{\text{it}} \cdot 2 \cdot q \cdot f \cdot A \cdot kT \cdot \ln \left(v_{\text{th}} \cdot n_i \cdot \sqrt{\sigma_n \cdot \sigma_p} \cdot \sqrt{t_{\text{em,e}} \cdot t_{\text{em,h}}} \right) \quad (2.19)$$

Auf der rechten Seite der Gleichung stehen neben der mittleren Grenzflächenzustandsdichte die Gatefläche A , die Temperatur T , die Pulsfrequenz f , die Elektronendriftgeschwindigkeit v_{th} , die intrinsische Ladungsträgerkonzentration von Silizium n_i und die Emissionszeiten t sowie Einfangsquerschnitte σ für Elektronen und Löcher.

Mit den Emissionszeiten für Elektronen und Löcher

$$t_{\text{em,e}} = \frac{|V_{\text{th}} - V_t|}{|\Delta V_g|} \cdot \frac{1 - \alpha}{f \cdot (1 + k_i)} \quad (2.20)$$

$$t_{\text{em,h}} = \frac{|V_{\text{th}} - V_t|}{|\Delta V_g|} \cdot \frac{\alpha}{f \cdot (1 + k_r)} \quad (2.21)$$

gilt damit für die gepumpte Ladung pro Puls:

$$Q_{\text{ss}} = \frac{I_{\text{CP}}}{f} = \bar{D}_{\text{it}} \cdot 2 \cdot q \cdot A \cdot kT \cdot \left\{ \ln \left(v_{\text{th}} \cdot n_i \cdot \sqrt{\sigma_n \cdot \sigma_p} \right) + \ln \left(\frac{|V_{\text{th}} - V_t|}{|\Delta V_g|} \cdot \frac{1}{f} \cdot \sqrt{\frac{\alpha \cdot (1 - \alpha)}{(1 + k_r) \cdot (1 + k_i)}} \right) \right\} \quad (2.22)$$

In einer Graphik von Q_{ss} gegenüber dem Logarithmus der Frequenz f ergibt sich eine Gerade mit der Steigung

$$\frac{dQ_{\text{ss}}}{d \log(f)} = \frac{2 \cdot q \cdot kT \cdot \bar{D}_{\text{it}}}{\log(e)} \cdot A = \frac{2 \cdot k[J/K] \cdot T[K]}{\log(e)} \cdot A [\text{cm}^2] \cdot \bar{D}_{\text{it}} [\text{eV}^{-1} \text{cm}^{-2}] \quad (2.23)$$

Dies entspricht demselben Zusammenhang, der von Groeseneken für einen Dreieckspuls abgeleitet wurde [2.15]. Eine einfache Bestimmung der mittleren Grenzflächenzustandsdichte \bar{D}_{it} ist somit möglich, wenn der Charge-Pumping-Strom bei verschiedenen Frequenzen f gemessen wird. Es wurde gezeigt, daß für diese Messung auch ein trapezförmiger Puls verwendet werden kann, wenn die Bedingungen in den Gleichungen 2.16 bis 2.18 erfüllt sind. Die Sensitivität der Methode ist in der Abbildung 2.5 dargestellt.

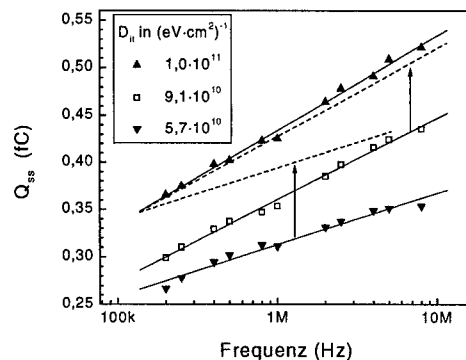


Abb. 2.5: Halblogarithmische Graphik der gepumpten Ladung pro Puls gegenüber der Pulsfrequenz. Für drei verschiedene Proben wurde die mittlere Grenzfächenzustandsdichte D_{it} bestimmt. Zusätzlich ist ein Vergleich der Steigungen der Geraden eingezeichnet.

2.2 Charakterisierung am MOS-Feldeffekttransistor

2.2.1 Einfluß der Gateoxidqualität auf wichtige Transistoreigenschaften

Mobile Ionen

Alkaliionen, insbesondere Na^+ , waren eine der Herausforderungen auf dem Weg zu einer stabilen MOS-Technologie. Unter dem Einfluß eines elektrischen Feldes driften diese Ionen in Siliziumdioxid, speziell bei einer erhöhten Temperatur. In extremen Fällen äußert sich das sogar bei Raumtemperatur durch Hystereseeffekte in der CV-Kurve. Durch mobile Ionen variiert die Einsatzspannung des Transistors, was zu einem instabilen Schaltungselement führt. Die Reinheit in der Halbleiterproduktion ist jedoch heute so hoch, daß Alkaliionen kein Qualitätskriterium mehr für Gateoxide sind. Die Prozeßverfahren berücksichtigen zudem die Bindung möglicher Alkaliionen, z.B. durch den Einbau von Cl^- mit Hilfe von HCl während der thermischen Oxidation.

Grenzflächenzustände und feste Oxidladungen

Ein fester Anteil an Ladungen im Gateoxid führt zu einer konstanten Verschiebung der Gatespannung um einen Betrag ΔV_g . Bei einer Verschiebung zu negativeren Gatespannungen ist der Ladungsbetrag insgesamt positiv und bei einer Verschiebung zu positiveren Gatespannungen insgesamt negativ. Die Ladung aufgrund von zusätzlich generierten

Grenzflächenzuständen führt also zu einer zusätzlichen Unbestimmtheit in der Einsatzspannung. In den Toleranzgrenzen für die Einsatzspannung müssen aber bereits Prozeßtoleranzen sowie die Temperaturabhängigkeit der Einsatzspannung berücksichtigt werden. Die Temperaturabhängigkeit der Einsatzspannung resultiert physikalisch aus der Temperaturabhängigkeit des Fermi-niveaus und beträgt etwa 1-2 mV/°C [2.16].

Ein weiterer Grund, speziell die Dichte der Grenzflächenzustände bzw. deren Generation zu minimieren, ist deren Einfluß auf die Steigung der Unterschwellgeraden („subthreshold slope“). In der Regel wird mit dem reziproken Wert, „subthreshold swing“ S , gearbeitet. Dieser Wert ist charakteristisch für das Schaltverhalten des Feldeffekttransistors. Er gibt an, um welchen Betrag die Gatespannung geändert werden muß, um den Drainstrom um eine Dekade zu erniedrigen.

$$S = \ln(10) \cdot \frac{\partial V_g}{\partial \ln(I_d)} \quad (2.24)$$

Der Wert S ist durch die Temperatur T , die Kapazität des Gateoxides C_{ox} und die Kapazität der Verarmungszone C_j bestimmt [2.16].

$$S = \frac{kT}{q} \cdot \ln(10) \cdot \left[1 + \frac{C_j}{C_{ox}} \right] \quad (2.25)$$

Addiert sich zur Kapazität der Verarmungszone die Kapazität der Grenzflächenzustände, dann erhöht sich S um den Betrag ΔS :

$$\Delta S = \frac{kT}{q} \cdot \ln(10) \cdot \frac{q \cdot D_{it}}{C_{ox}} \quad (2.26)$$

Der Drainstrom bei einer Gatespannung $V_g = 0$ wird als Off-Strom I_{off} bezeichnet und bestimmt die Verlustleistung der Schaltung im Ruhezustand. Der Drainstrom I_i bei $V_g = V_i$ ist für typische NMOSFETs in der Größenordnung von 10^{-7} A/μm. Wird für den Off-Strom eine bestimmte Zielvorgabe gemacht, dann definiert die Steigung der Unterschwellgeraden eine minimal mögliche Einsatzspannung $V_{i,min}$:

$$V_{i,min} = S \cdot \log \left(\frac{I_i}{I_{off}} \right) \quad (2.27)$$

Der Skalierung ist damit durch die Temperatur ein natürliches Limit gesetzt, denn der theoretische Minimalwert für S bei 300 K und Silizium ist etwa 60 mV/Dekade [2.16]. Grenzflächenzustände erhöhen S und damit das Skalierungslimit. Eine Verkleinerung der Grenzflächenzustandsdichte wird durch einen Temperaturschritt in Formiergas nach der Metallisierung erreicht.

Eine wichtige Größe im Zusammenhang mit den Stromeigenschaften eines Feldeffekttransistors ist die Beweglichkeit der freien Ladungsträger im Inversionskanal. Es wird zwischen der Feldeffektbeweglichkeit μ_{FE} und der effektiven Beweglichkeit μ_{eff} unterschieden. Die jeweiligen Definitionen der Beweglichkeiten sowie der Zusammenhang

mit der Transferleitfähigkeit g_m werden im Anhang A behandelt. Zur Beschreibung des Drainstromes sollte nur die effektive Beweglichkeit verwendet werden, weil bei der Feldeffektbeweglichkeit Unterschiede bezüglich Gateoxid und Ladungsträger überschätzt werden können [2.16]. Die Feldeffektbeweglichkeit ist jedoch ein guter Parameter zur Messung der Degradation eines Transistors, da zum einen ihr Maximalwert gleich dem der effektiven Beweglichkeit ist [2.16] und zum anderen ein direkter Zusammenhang zur leicht meßbaren Transferleitfähigkeit g_m besteht. Wird die effektive Beweglichkeit μ_{eff} nicht gegenüber dem vertikalen elektrischen Feld aufgetragen, sondern gegenüber einem effektiven elektrischen Feld, dann zeigt sich unabhängig von der Substratdotierung und der Substratspannung ein universeller Verlauf [2.17]. Die Definition des effektiven elektrischen Feldes ist im Anhang A aufgeführt. Dieser universelle Verlauf wird durch drei Streumechanismen für die Ladungsträger im Inversionskanal geprägt. Eine schematische Darstellung ist in der Abbildung 2.6 gezeigt [2.18]. Der Anteil der Coulombstreuung ist bei kleinen effektiven Feldern dominant und wird durch drei Arten von Streuzentren verursacht [2.18]:

- Fremdatome im Silizium (Substratdotierung),
- Ladungen in Grenzflächenzuständen und
- feste Oxidladungen.

Die Zunahme an Oxidladungen und Grenzflächenzuständen durch elektrischen Streß führt zu einer Degradation des Transistors. Die Transferleitfähigkeit g_m ist dabei einer der degradierenden Transistorparameter.

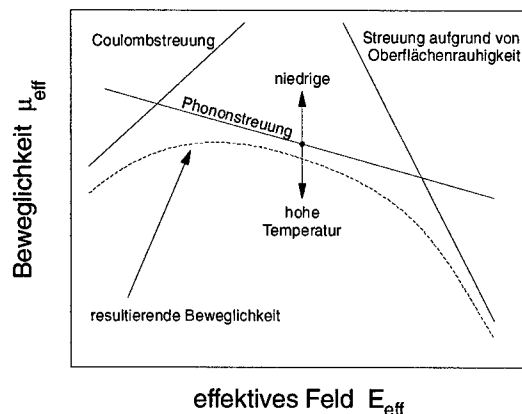


Abb. 2.6: Schematische Darstellung der effektiven Beweglichkeit in Abhängigkeit vom effektiven elektrischen Feld und vom jeweils dominanten Streumechanismus. Die durch Phononstreuung bestimmte Beweglichkeit ist temperaturabhängig. Die insgesamt resultierende Beweglichkeit ist gestrichelt eingezeichnet (Graphik gemäß [2.18]).

2.2.2 „Hot Carrier Stress“

Ladungsträgerinjektion in das Gateoxid

Die Potentialbarriere für die Injektion von Elektronen vom Leitungsband des Siliziums in das Leitungsband von Siliziumdioxid beträgt ca. 3,1 eV [2.16]. Die Bandlücke von Siliziumdioxid ist etwa 9 eV [2.2], so daß die Potentialbarriere für die Injektion von Löchern bezogen auf die jeweiligen Valenzbandkanten ca. 4,8 eV beträgt. Aus diesem Grund wird oft nur die Injektion von Elektronen betrachtet. Es existieren sechs Mechanismen für die Injektion von Ladungsträgern [2.19]:

- 1) Fowler-Nordheim-Tunneln
- 2) Direktes Tunneln
- 3) SHE (Substrate Hot-Electron Injection“), SHI („Substrate Hot-Hole Injection“)
- 4) CHE („Channel Hot-Electron Injection“)
- 5) DAHC („Drain Avalanche Hot-Carrier Injection“)
- 6) SGHE („Secondarily Generated Hot-Electron Injection“)

Fowler-Nordheim-Tunneln und direktes Tunneln werden im Kapitel 4 behandelt. Im Fall von SHE und SHI überwinden sehr energiereiche Ladungsträger die Potentialbarriere ohne quantenmechanisches Tunneln. Dies können z.B. thermisch generierte Elektronen oder durch Strahlung generierte Ladungsträger sein. Eine gezielte Injektion ist möglich, wenn aus einem in Vorwärtsrichtung gepolten p-n-Übergang im Substrat die Ladungsträger in Richtung des Gateoxides beschleunigt werden. SHI ist also für einen PMOSFET mit n-Wanne und p-Substrat möglich. Die meisten Löcher werden zwar von Source und Drain aufgenommen, aber die energiereichsten können in das Gateoxid injiziert werden [2.20]. In den Mechanismen 1) bis 3) ist nur der eindimensionale Transport von Elektronen senkrecht zur Grenzfläche zwischen Oxid und Siliziumsubstrat involviert.

Die Degradation durch heiße Ladungsträger entsteht im MOS-Feldeffekttransistor jedoch durch den Ladungstransport im Inversionskanal und ist ein zweidimensionales Problem. Hierzu tragen die Mechanismen 4) bis 6) bei. Die Schädigung des Gateoxides durch injizierte Ladungen ist in Drainnähe lokalisiert, da dort die Inversionsladungen ihre maximale Beschleunigung im lateralen elektrischen Feld zwischen Source und Drain erreichen. Im Fall von CHE erlangen einige Elektronen im Inversionskanal genügend Energie, um die Potentialbarriere zum Oxid zu überwinden, bevor sie einem inelastischen Stoß unterliegen („lucky electron model“ [2.21, 2.22, 2.23]). DAHC entsteht durch Stoßionisation in Drainnähe und führt zur Injektion von Elektronen und Löchern. SGHE basiert auf Minoritätsladungsträger, die durch sekundäre Stoßionisation oder Bremsstrahlung entstanden sind und anschließend analog zu SHE injiziert werden.

Aufgrund der Stoßprozesse und der sekundären Generationsprozesse verursachen energiereiche Ladungsträger neben einer zusätzlichen Komponente zum Gatestrom auch

einen Substratstrom. Dieser ist beim NMOSFET wegen Avalanchemultiplikation wesentlich größer als der Gatestrom.

„Hot Carrier Stress“

Für die Charakterisierung der Stabilität eines MOS-Feldeffekttransistors gegen heiße Ladungsträger wird dieser unter den Bedingungen der maximalen Degradation betrieben. Die maximale Degradation ist abhängig von den Streßbedingungen und vom Kanaltyp [2.24, 2.25]. Für den NMOSFET ist das beim Substratstrommaximum im Bereich $V_d/3 < V_g < V_d/2$ der Fall. V_d ist die Drainspannung und V_g die Gatespannung. In diesem Spannungsbereich findet die Generation von Grenzflächenzuständen durch die Injektion heißer Elektronen und heißer Löcher in das Gateoxid statt. Für den PMOSFET liegt die maximale Degradation bei $V_g \approx V_d$ vor. Dafür trägt im wesentlichen das Haften von injizierten Elektronen und die Generation von Grenzflächenzuständen bei.

Die Charakterisierung zielt auf die Bestimmung einer Lebensdauer ab, innerhalb der die Änderung wichtiger Transistorparameter ein gewisses Maß nicht überschreitet. Dies sind z.B. die Änderung der Einsatzspannung V_i oder die Änderung des Maximums der Transferleitfähigkeit g_m . Als Grenzen der Degradation werden z.B. festgelegt:

$$\frac{\Delta g_m}{g_m} < 10\%, \quad \Delta V_i < 50\text{mV} \quad (2.28)$$

Die im beschleunigten Test gemessene Zeit muß auf Betriebsbedingungen extrapoliert werden, um eine Lebensdauer des untersuchten Systems zu erhalten. Für den NMOSFET ist die Degradation der Transistorparameter eine Potenzfunktion der Streßzeit [2.19]:

$$\Delta Y(t) = A \cdot t^b \quad (2.29)$$

Ein Beispiel für einen NMOSFET ist in der Abbildung 2.7 gezeigt. Der Parameter b hängt stark von der Gatespannung ab, hat jedoch nur eine kleine Abhängigkeit von der Drainspannung. Dies läßt darauf schließen, daß sich die Steigung b in der doppellogarithmischen Darstellung mit dem Mechanismus der Hot-Carrier-Injektion ändert [2.19]. Der Parameter A ist ein Maß für die Höhe der Degradation und hängt stark von der Drainspannung ab, wenig hingegen von der Gatespannung. Es zeigt sich empirisch [2.19]:

$$A \propto \exp\left(-\frac{\alpha}{V_d}\right) \quad (2.30)$$

Für die Extrapolation der Lebensdauer τ läßt sich verwenden [2.19, 2.26]:

$$\tau \propto \exp\left(\frac{\alpha}{b \cdot V_d}\right) \quad (2.31)$$

$$\tau \cdot \frac{I_d}{W} \propto \left[\frac{I_{\text{sub}}^{\text{max}}}{I_d} \right]^{\frac{\phi_{it}}{\phi_i}} \quad \text{mit} \quad \frac{\phi_{it}}{\phi_i} \approx 3 \quad (2.32)$$

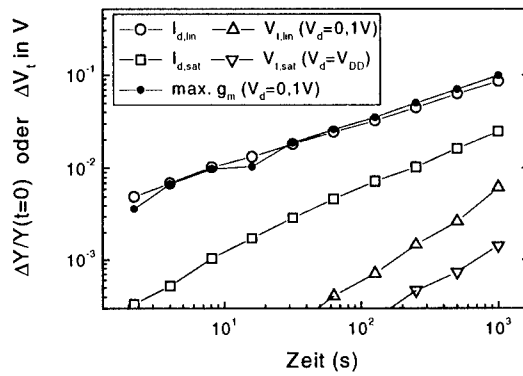


Abb. 2.7: Beispiel für die Degradation eines NMOSFETs durch „Hot Carrier Stress“. Y steht für einen degradierenden Transistorparameter.

Der Quotient aus maximalem Substratstrom und Drainstrom in der Gleichung 2.32 wird als Multiplikationsfaktor bezeichnet und ist ein Maß für die Elektron-Loch-Paarerzeugung in Drainnähe. Die Größen ϕ_{it} und ϕ_i sind nach Hu [2.26] die kritische Energie eines Elektrons, um einen Grenzflächenzustand zu erzeugen, bzw. die minimale Energie eines Elektrons, um Stoßionisation zu verursachen.

Für die Degradation des PMOSFETs ergibt sich eine logarithmische Abhängigkeit von der Zeit [2.27, 2.28, 2.29]:

$$\Delta Y = S \cdot \ln\left(\frac{t}{\Theta}\right) \quad (2.33)$$

$$\ln(\Theta) \propto \frac{1}{V_d - V_{d,sat}} \quad (2.34)$$

$$S \propto \frac{1}{V_d - V_{d,sat}}$$

$V_{d,sat}$ ist die Sättigungsdrainspannung und Θ ist ein Maß für den Einfang an Elektronen in Haftstellen.

Charakterisierung des Gateoxides

Die Degradation durch heiße Ladungsträger ist abhängig von der Technologie des Feldeffekttransistors. Zur Minimierung der Degradation wurde z.B. ein niedrig dotierter p-n-Übergang (LDD: „lightly doped drain“) eingeführt, um einen kleineren Potentialgradienten auf der Drainseite zu erhalten. Sollen zwei Gateoxide bezüglich ihrer Stabilität gegenüber

heiße Ladungsträger miteinander verglichen werden, muß die Unabhängigkeit von diesen Einflüssen gegeben sein. Somit ist beim „Hot Carrier Stress“ bzw. bei der Auswertung darauf zu achten, daß

- a) die MOSFET-Technologie,
- b) das vertikale elektrische Feld an der Oberfläche des Siliziumsubstrates und
- c) die Drainspannung während des Streßexperimentes

jeweils gleich sind.

2.2.3 „Bias Temperature Stress“

Eine elektrische Belastung bei erhöhter Temperatur wurde früher verwendet, um den Effekt der mobilen Ionenladungen im Gateoxid zu untersuchen. Seit der Lösung dieses Problems treten unter dem Einfluß eines elektrischen Feldes die Generierung von Haftstellen und Oxidladungen in den Vordergrund. Die Verschiebung der Einsatzspannung, speziell bei einem Streß mit negativer Gatespannung, wurde sehr früh als Degradationsproblem erkannt [2.30]. Bei positiven Gatespannungen ist diese Instabilität in der Einsatzspannung auf die Erzeugung von donatorartigen Grenzflächenzuständen zurückzuführen [2.31]. Im Fall von negativen Gatespannungen trägt dazu sowohl die Generation von Grenzflächenzuständen als auch die von positiven Oxidladungen bei. Die Testmethode wird als „Bias Temperature Stress“ (BTS), bzw. für negative Spannungen speziell als „Negative Bias Temperature Stress“ (NBTS) bezeichnet. Im Gegensatz zum „Hot Carrier Stress“ (HCS) wird der BTS homogen durchgeführt, d.h. Source und Drain liegen am gleichen Potential. Die Degradation ist ebenso wie bei HCS eine Potenzfunktion der Streßzeit:

$$\Delta Y(t) = C \cdot t^d \quad (2.35)$$

Neben der Einsatzspannung degradieren auch andere Transistorparameter. In der Abbildung 2.8 ist z.B. der Drainstrom gezeigt. Der Exponent d der Potenzfunktion ist jeweils in etwa gleich.

Die Publikationen befassen sich zunehmend mit diesem Degradationsphänomen [2.32, 2.33, 2.34, 2.35, 2.36], da die Instabilität des PMOSFETs bei zukünftigen Technologiegenerationen immer mehr an Bedeutung gewinnt. Die Ursache für die Degradation ist noch nicht genau geklärt.

Sollen unterschiedliche Gateoxide hinsichtlich ihrer Stabilität gegenüber BTS untersucht werden, ist auf eine Vergleichbarkeit zu achten. Im Gegensatz zum „Hot Carrier Stress“ ist dies einfacher, da ein homogener Streß vorliegt. Die Verwendung derselben Transistortechnologie ist unabdingbar. Neben der Temperatur ist auf das vertikale elektrische Feld zu achten.

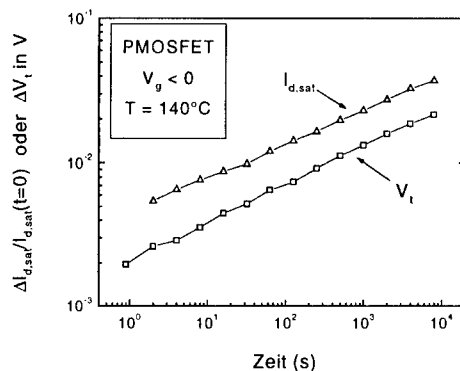


Abb. 2.8: Beispiel für die Degradation eines PMOS-Kurzkanaltransistors während „Bias Temperature Stress“ bei negativer Gatespannung und 140°C.

2.3 Zuverlässigkeitsuntersuchungen

Es ist bekannt, daß ein dielektrischer Durchbruch des Gateoxides, der während der Betriebsdauer entsteht, eine der hauptsächlichen Fehlermechanismen in integrierten Schaltkreisen mit MOS-Technologie ist [2.37]. Dieser Durchbruch wird als „Time Dependent Dielectric Breakdown“ (TDDB) bezeichnet. Ein wesentlicher Teil dieser Arbeit beschäftigt sich mit neuen Degradationsphänomenen ultradünner Gateoxide und greift auf grundlegende Begriffe und Techniken der Zuverlässigkeitsanalysen von Dielektrika zurück. Daher werden diese in den folgenden Kapiteln etwas ausführlicher behandelt.

2.3.1 Elektrische Streßmethoden

Zur Charakterisierung der Zuverlässigkeit von Dielektrika werden Streßmessungen an Kapazitätsstrukturen oder parallel geschalteten Transistoren durchgeführt. Oxide sind keine idealen Isolatoren und unter dem Einfluß eines elektrischen Feldes werden Leckströme gemessen. Eine Auflistung von Leckstrommechanismen ist z.B. in [2.2] enthalten. An einer MOS-Kapazität kann entweder ein Strom oder eine Spannung eingepreßt werden und es wird zwischen zwei grundlegenden Streßmethoden unterschieden:

- „Time Dependent Dielectric Breakdown“-Test (TDDB-Test)
- E-Rampen-Test für sehr schnelle Analysen.

Bei der TDDB-Testmethode wird unter konstanten Streßbedingungen die Zeit bis zum dielektrischen Durchbruch bestimmt. Bei der E-Rampen-Methode wird eine Strom-Spannungs-Rampe durchfahren. Die charakteristische Größe ist das elektrische Feld, bei dem der dielektrische Durchbruch erfolgt ist.

Die Durchbruchereignisse werden beim E-Rampen-Test in drei Klassen eingeteilt. Ein sofortiger Durchbruch bei sehr kleinen elektrischen Feldern wird mit „A-mode“ bezeichnet. Eine enge Verteilung in der Nähe von ca. 12 MV/cm wird mit „C-mode“ und Durchbrüche bei kleineren Feldstärken mit „B-mode“ bezeichnet. Durchbrüche der Klasse „A-mode“ und „B-mode“ werden durch extrinsische Fehler verursacht, die auf Prozeßeinflüsse zurückzuführen sind. Im TDDB-Test weisen sie kurze Streßzeiten bis zum Durchbruch auf. Die Klasse „C-mode“ zeigt im TDDB-Test eine enge Verteilung bei einer hohen Streßzeit. Diese wird als intrinsische Verteilung bezeichnet. Die Bezeichnung ist historisch bedingt. Es handelt sich dabei nicht um eine intrinsische Eigenschaft von Siliziumdioxid.

Für Siliziumdioxid hat sich die Ladung, welche bis zum dielektrischen Durchbruch das Oxid durchquert hat, als charakteristisch erwiesen. Dieser Wert wird mit Q_{BD} bezeichnet und berechnet sich aus der Zeit bis zum Durchbruch t_{BD} und der injizierten Stromdichte J_{inj} :

$$Q_{BD} = \int_0^{t_{BD}} J_{inj}(t) dt \quad (2.36)$$

Der Q_{BD} -Wert liegt in der Größenordnung von 1-10 C/cm². Mit der Skalierung der Gateoxidicke ab der 0,18 µm-Generation wird diese Richtgröße in zunehmendem Maße durch den direkten Tunnelstrom verfälscht.

2.3.2 Beschleunigungsmodelle

Alle Zuverlässigkeitstests sind unter beschleunigten Bedingungen durchzuführen, um kurze Meßzeiten zu erhalten. Anhand von Modellen muß von einer kurzen Zeit bis zum Auftreten der Fehlfunktion auf die reale Lebensdauer unter Betriebsbedingungen extrapoliert werden. Die Beschleunigung des dielektrischen Durchbruches erfolgt über die Temperatur und das elektrische Feld. Es existieren zwei Modelle, das E-Modell und das 1/E-Modell, die seit vielen Jahren in Konkurrenz zueinander stehen.

E-Modell

Die Zeit bis zum Durchbruch wurde als erstes erfolgreich mit der folgenden Relation beschrieben:

$$\ln(t_{BD}) \propto \frac{\Delta H_0}{kT} - \gamma \cdot E_{ox} \quad (2.37)$$

ΔH_0 ist die Enthalpie für einen dielektrischen Durchbruch in SiO₂ und wird gewöhnlich als Aktivierungsenergie E_a bezeichnet. E_{ox} ist das elektrische Feld im Oxid und γ ist der Feldbeschleunigungsfaktor. Diese Relation wurde zunächst als empirische eingeführt [2.38, 2.39, 2.40], später aber von McPherson et al. thermochemisch begründet [2.41, 2.42, 2.43,

2.44, 2.45, 2.46]. Die Beschleunigung der Durchbruchzeiten bezogen auf zwei elektrische Felder E_1 und E_2 wird somit im allgemeinen wie folgt dargestellt:

$$t_1 = t_2 \cdot \exp[-\gamma \cdot (E_1 - E_2)] \quad (2.38)$$

Der Feldbeschleunigungsfaktor γ ist entsprechend aus experimentellen Daten ermittelbar:

$$\gamma = \frac{\ln\left(\frac{t_1}{t_2}\right)}{(E_2 - E_1)} \quad (2.39)$$

Bei der Umrechnung zwischen zwei verschiedenen Temperaturen wäre nach Gleichung 2.37 einfach ein multiplikativer Faktor analog zum Arrhenius-Modell zu berücksichtigen:

$$\exp\left[\frac{E_a}{k} \left(\frac{1}{T_1} - \frac{1}{T_2}\right)\right] \quad (2.40)$$

Es wird jedoch sowohl eine Temperaturabhängigkeit des Feldbeschleunigungsfaktors als auch eine Abhängigkeit der Aktivierungsenergie vom elektrischen Feld festgestellt [2.40, 2.41, 2.47]. Damit ist eine exakte Projektion der Lebensdauer nicht durch eine multiplikative Verknüpfung der Feldbeschleunigung und der Temperaturbeschleunigung möglich.

Das E-Modell wird auch als thermodynamisches Modell oder thermochemisches Modell bezeichnet.

1/E-Modell

Das 1/E-Modell setzt den dielektrischen Durchbruch in Zusammenhang mit dem Haften von Löchern im Oxid („hole trapping“). Dieser Vorgang ist an den Elektronenstrom durch das Oxid gekoppelt. Die lokalen Veränderungen des elektrischen Feldes durch Löcher erhöhen den Elektronenstrom und damit wieder die Anzahl von Löchern im Oxid. Diese positive Rückkopplung führt zum dielektrischen Durchbruch und wird als „Hole-Induced Breakdown Model“ bezeichnet. Zunächst wurde vermutet, daß die Löcher im Oxid durch Stoßionisation von injizierten Elektronen entstehen [2.48, 2.49]. Dieser Mechanismus wurde aber später verworfen und durch das sogenannte „Anode Hole Injection Model“ ersetzt [2.50, 2.51, 2.52]. Elektronen, die von der Kathode aus in das Oxid injiziert wurden, verlieren einen Großteil ihrer Energie im Leitungsband der Anode. Diese Energie wird an Elektronen im Valenzband der Anode abgegeben, die somit bis zur Leitungsbandkante angehoben werden. Im Valenzband verbleiben heiße, also energiereiche Löcher, die gemäß Fowler-Nordheim zur Kathode zurücktunneln. Im Oxid generieren diese Löcher teilweise neue Haftstellen und werden auch selbst in Haftstellen eingefangen. Da sowohl der Fowler-Nordheim-Tunnelstrom als auch die Generationsrate für Löcher proportional zu $1/E$ sind, wurde eine 1/E-Abhängigkeit der Zeit bis zum Durchbruch vermutet. Die ersten experimentellen Daten mit einer 1/E-Abhängigkeit sind in [2.48] gezeigt und das 1/E-Modell wurde in [2.53] erstmals ausführlicher formuliert. Für die Zeit bis zum Durchbruch t_{BD} ergibt sich:

$$t_{BD} \propto \exp\left(\frac{G}{E_{ox}}\right) \quad (2.41)$$

Der Koeffizient G ist die Summe aus einem Anteil B für den Fowler-Nordheim-Tunnelstrom und einem Anteil H für die Generation von Löchern. Für Oxide dicker ca. 12 nm ist $G \approx 350$ MV/cm, für dünnere Oxide wird G größer [2.50].

Ein extrinsischer Oxiddefekt wird im 1/E-Modell als eine Dünnstelle im Oxid betrachtet. Das Oxid besitzt an dieser Stelle also nur die effektive Oxiddicke $x_{eff} = x_{ox} - \Delta x_{ox}$ („Effective Oxide Thinning Model“). Damit ergibt sich:

$$t_{BD} = \tau_0 \cdot \exp\left[\left(1 - \frac{\Delta x_{ox}}{x_{ox}}\right) \cdot \frac{G}{E_{ox}}\right] \quad (2.42)$$

Vergleich von E-Modell und 1/E-Modell – Neue Zuverlässigkeitsmodelle

Das E-Modell ist konservativer als das 1/E-Modell, da es kleinere extrapolierte Zeiten ergibt. In der Abbildung 2.9 sind Daten aus verschiedenen Arbeiten zusammengestellt. Für übliche Streßzeiten zwischen 1 s und 10^4 s ergibt sich immer das Problem, daß sich beide Modelle an die Meßdaten gut anpassen. Eine deutliche Abweichung vom linearen E-Verhalten wurde bisher nur für sehr kleine Streßzeiten, also bei sehr hohen elektrischen Feldstärken gemessen. Obwohl in letzter Zeit Langzeitversuche unternommen wurden, um eine experimentelle Überprüfung zu erreichen [2.56, 2.57, 2.58], ist die Modellfrage noch nicht geklärt.

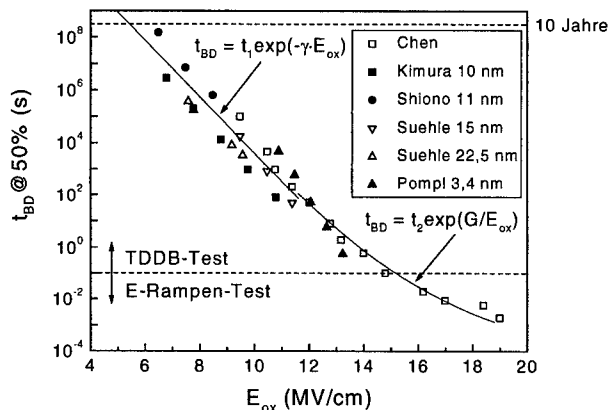


Abb. 2.9: Experimentelle Daten für dicke Oxide [2.54, 2.55, 2.56, 2.59] und ein 3,4 nm dickes Oxid. Die Anpassungen für die jeweiligen Modelle sind aus [2.59] übernommen. Die Daten entsprechen einem Test mit einer aktiven Fläche von $0,001 \text{ cm}^2$ und einer Temperatur von 125°C .

Der Grund für die sehr hohen extrapolierten Zeiten im 1/E-Modell ist der im „Anode Hole Injection Model“ vorgeschlagene Mechanismus. Eine kleine Feldstärke bedeutet eine kleine Potentialdifferenz zwischen Kathode und Anode. Damit wird auch die maximal mögliche potentielle Energie eines generierten Loches klein. Da der Fowler-Nordheim-Tunnelstrom exponentiell von der Höhe der Potentialbarriere abhängt, sinkt damit die Tunnelwahrscheinlichkeit sehr stark. Es werden zur Zeit neue Modelle entwickelt, die z.B. sowohl das E-Modell als auch das 1/E-Modell für verschiedene Spannungsbereiche berücksichtigen [2.59, 2.60]. Wie in der Abbildung 2.9 gezeigt, wird darin die entscheidende Extrapolation zu kleinen elektrischen Feldern nach dem E-Modell durchgeführt. Dabei ist zu bemerken, daß das „Unified Gate Oxide Reliability Model“ in [2.59] von der Gruppe um Chenming Hu von der University of California, Berkeley, vorgeschlagen wurde, welche das 1/E-Modell hervorgebracht und ausformuliert hat.

2.3.3 Statistische Datenauswertung

Die Qualität des Oxides ist nicht durch die intrinsische Verteilung bestimmt, sondern durch einen kleinen Anteil extrinsischer Fehler. Daher wird in der Zuverlässigkeitsanalytik mit der Weibull-Verteilung eine Extremwertverteilung zur Datenauswertung herangezogen. Im Anhang B ist eine allgemeine Formulierung der Weibull-Verteilung angegeben. Da die statistische Größe die Zeit t ist, wird die zweiparametrische Weibull-Verteilung verwendet. Der Formparameter ist im folgenden mit β bezeichnet. Der Maßstabsparameter a entspricht einer charakteristischen Lebensdauer, z.B. der Zeit $t_{63\%}$ bei welcher 63% der Proben fehlerhaft sind. Für die Verteilungsfunktion $F(t)$ gilt:

$$F(t) = 1 - \exp \left[- \left(\frac{t}{a} \right)^\beta \right] \quad (t > 0) \quad (2.43)$$

Die Verteilungsfunktion $F(t)$ ist der Anteil der zur Zeit t bereits defekten Teststrukturen und wird oft als Ausfallanteil oder Fehleranteil bezeichnet. Zur Analyse der Daten wird folgende Beziehung herangezogen:

$$\ln[-\ln(1 - F(t))] = \beta \cdot \ln(t) - \beta \cdot \ln(a) \quad (2.44)$$

Daher wird nicht die kumulative Verteilung $F(t)$, sondern der linke Teil der Gleichung 2.44 gegen $\log(t)$ aufgetragen. Dieser Plot wird als Weibull-Plot oder Gumbel-Plot und die Steigung β der Geraden als „Weibull slope“ bezeichnet. Ein Beispiel ist in der Abbildung 2.10 gezeigt.

Problem der 100% Ausfälle

Für die Probe mit der längsten Streßzeit t_{\max} ergibt sich das Problem, daß $F(t_{\max}) = 1$ ist. Dieser Wert ist in der Weibull-Statistik nicht definiert (siehe Gleichung 2.44). In der Praxis wird daher dieser Wert korrigiert, so daß $F(t_{\max}) < 1$ ist. Außerdem stellt eine Messung immer eine Stichprobe mit endlicher Anzahl dar. Die 100% Ausfälle sind zwar in der Messung bei t_{\max} erreicht, aber in Wirklichkeit könnte es ja doch noch höhere Zeiten geben. Aus diesem Grund krümmt sich die Verteilung im Weibull-Plot bei hohen F konvex, abweichend von der

exakten intrinsischen Geraden, die zu beliebig kleinen Fehleranteilen bzw. unendlich großen Testzahlen führt.

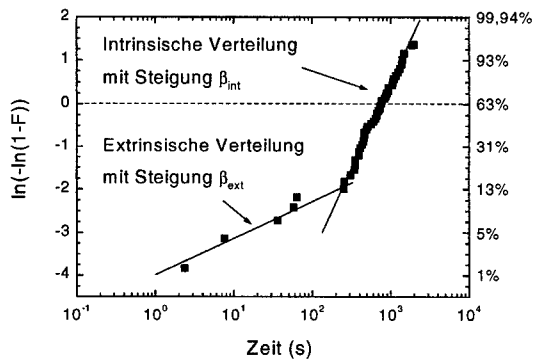


Abb. 2.10: Beispiel für einen Weibull-Plot. Die TDDB-Verteilung zeigt den intrinsischen Ast mit der Steigung β_{int} und einen extrinsischen Ast mit der Steigung β_{ext} .

Bimodale Verteilung

In der Abbildung 2.10 ist zu erkennen, daß die intrinsische Verteilung eine größere Steigung als die extrinsische Verteilung aufweist. Dies ist in Übereinstimmung mit den folgenden Fällen, die bei der Weibull-Verteilung unterschieden werden können [2.61]:

- $\beta < 1$: $\lambda(t)$ nimmt mit zunehmender Zeit ab und die Weibull-Verteilung wird zur Charakterisierung der frühen Fehlerphase herangezogen (extrinsische Fehler).
- $\beta > 1$: $\lambda(t)$ nimmt mit zunehmender Zeit zu und die Weibull-Verteilung wird zur Charakterisierung der späten Fehlerphase herangezogen (intrinsischer Fehler).

Für die bimodale Weibull-Verteilung gilt:

$$\begin{aligned}
 Y_{\text{int}} &= 1 - F_{\text{int}}(t) = \exp \left[- \left(\frac{t}{t_{63\%}} \right)^{\beta_{\text{int}}} \right] \\
 Y_{\text{ext}} &= 1 - F_{\text{ext}}(t) = \exp \left[- \left(\frac{t}{t_{63\%}} \right)^{\beta_{\text{ext}}} \right] \\
 Y_{\text{total}} &= 1 - F(t) = Y_{\text{int}} \cdot Y_{\text{ext}}
 \end{aligned} \tag{2.45}$$

2.3.4 Extrapolation der Lebensdauer auf Betriebsbedingungen

Die Zielvorgaben für die Qualifikation eines Gateoxides sind die minimale Lebensdauer und der maximale Fehleranteil in dieser Zeit. Typisch für eine geforderte Lebensdauer ist 10 Jahre. Der Fehleranteil wird z.B. in ppm („parts per million“) angegeben. Beide Angaben sind in der „failure unit“ FIT verarbeitet:

$$1 \text{ FIT} = 1 \text{ Fehler}/10^9 \text{ Stunden} \approx 0,01\% \text{ Fehler}/10 \text{ Jahre} \quad (2.46)$$

Bei der Extrapolation der gemessenen Durchbruchzeiten müssen verschiedene Parameter des Produktes mit einbezogen werden. Dies sind

- die Betriebsspannung, woraus sich das elektrische Feld im Oxid E_{ox} ergibt,
- die Temperatur im Betrieb,
- die gesamte aktive Gatefläche pro Chip,
- der „duty cycle“ bzw. der „duty factor“ und
- ein möglicher „burn-in“.

Die Extrapolation des elektrischen Feldes und der Spannung betrifft die Modelldiskussion im Kapitel 2.3.2. Ein „burn-in“ kann nur bei redundant ausgelegten Schaltungen, wie z.B. Speicherchips, verwendet werden. Er eliminiert die frühe Ausfallphase. Der „duty cycle“ legt den Anteil an der Betriebsdauer fest, in dem das Oxid dem elektrischen Streß bei Betriebsspannung ausgesetzt ist. Ein „duty factor“ f_{duty} von 100 bedeutet, daß das Oxid während einem Prozent der Betriebsdauer t_{use} belastet wird.

$$t_{use} = t_{active} \cdot f_{duty} \quad (2.47)$$

Sowohl der extrinsische Durchbruch als auch der intrinsische Durchbruch erfolgen an der schwächsten Stelle im Oxid („weakest-link character“). Dadurch wird die Fehlerverteilung abhängig von der Fläche und die dazugehörige Statistik ist die Poisson-Statistik (siehe Anhang B). Es kann unter Verwendung von Poisson-Statistik und Weibull-Verteilung gezeigt werden, daß die Zeit bis zum dielektrischen Durchbruch t_{BD} einem Potenzgesetz der Testfläche A folgt und der Exponent gleich dem inversen Weibull-Parameter β ist.

$$t_{BD} \propto A^{-n} = A^{-\frac{1}{\beta}} \quad (2.48)$$

Eine genaue experimentelle Untersuchung und mehrere Literaturstellen sind in [2.62] enthalten. Damit ergibt sich im Weibull-Plot eine vertikale Verschiebung der Verteilungen, die zu Kapazitäten mit verschiedenen Flächen $A_1 > A_2$ gehören. Der additive Betrag ist:

$$\Delta[\ln(-\ln(1-F))] = \ln\left(\frac{A_1}{A_2}\right) \quad (2.49)$$

Diese Verschiebung ergibt wegen der steilen intrinsischen Geraden keine großen Veränderungen für die intrinsische Durchbruchzeit. Die Gerade der extrinsischen Durchbrüche ist aber kleiner als eins. Hier wirkt sich die Extrapolation der kleinen Testfläche auf die große, gesamte aktive Gateoxidfläche im Produkt sehr stark aus.

3 Nitridierte Gateoxide – Stickstoffbarriere gegen Bordiffusion

3.1 „Dual Workfunction CMOS“-Logik

Die CMOS-Logik („Complementary MOS“) reduziert die Gleichstromleistungsaufnahme von MOS-Inverter durch die Verknüpfung von NMOSFET und PMOSFET. Der Betrag der Einsatzspannungen muß dabei vergleichbar sein: $V_t(\text{PMOSFET}) \approx -V_t(\text{NMOSFET})$. Bei der Verwendung identischer Gatematerialien für den NMOSFET und den PMOSFET ergibt sich das Problem unterschiedlicher Austrittsarbeitsdifferenzen Φ_{ms} zwischen Gate und Wanne. Der Zusammenhang zwischen Austrittsarbeitsdifferenz und Einsatzspannung ist aus den folgenden Gleichungen ersichtlich.

$$V_t = V_{th} + 2 \cdot \Psi_B + \frac{1}{C_{ox}} \sqrt{2 \cdot \epsilon_{Si} \cdot q \cdot N_{sub} \cdot (2 \cdot \Psi_B + V_{sub})} \quad (3.1)$$

$$V_{th} = \Phi_{ms} - \frac{Q_f}{C_{ox}} \quad (3.2)$$

Austrittsarbeitsdifferenzen für verschiedene Gatematerialien und in Abhängigkeit von der Dotierung des Siliziumsubstrates sind in der Abbildung 3.1 gezeigt. Eine entartet dotierte Polysiliziumschicht konnte zunächst nur durch einen LPCVD-Prozeß realisiert werden. Daher waren in CMOS-Technologien die Gates von NMOSFET und PMOSFET jeweils n^+ -dotiert. Der Unterschied in der Austrittsarbeitsdifferenz beträgt in diesem Fall 0,5-0,9 V (siehe Abbildung 3.1). Das Problem wurde durch eine oberflächennahe Gegendotierung des Kanals des PMOSFETs gelöst. Für die Justierung der Einsatzspannungen von NMOSFET und PMOSFET verwendete man also eine Kanalimplantation vom p-Typ (Akzeptoren). Der PMOSFET wird dadurch zu einem Transistor mit vergrabenem Kanal („Buried Channel p-FET“). Eine Beschreibung des „Buried Channel p-FET“ und des Unterschiedes zum „Surface Channel p-FET“ ist im Anhang A enthalten. Unter Berücksichtigung der zusätzlichen Kapazität der Raumladung des verarmten „Buried Channel“-Gebietes C_{bc} (siehe Anhang A) kann die Einsatzspannung des PMOSFETs symmetrisch zum NMOSFET justiert werden.

Durch die zusätzliche Kapazität C_{bc} ist die Ansteuerung bei Transistoren mit vergrabenem Kanal deutlich schlechter als bei Transistoren mit Oberflächenkanal. Die Steigung der Unterschwelliggeraden ist kleiner, bzw. S ist größer:

$$S = \frac{kT}{q} \cdot \ln(10) \cdot \left[1 + C_D \cdot \left(\frac{1}{C_{ox}} + \frac{1}{C_{bc}} \right) \right] \quad (3.3)$$

Im Vergleich zum FET mit Oberflächenkanal ist der wesentliche Nachteil des „Buried Channel“-FET durch die schlechteren Kurzkanaleigenschaften gegeben [3.1, 3.2]. Mit Hochstromimplantationsanlagen ist es mittlerweile z.B. möglich, Polysiliziumgates entartet zu dotieren. Dies wurde ab der 0,35µm-Logikgeneration genutzt, um unterschiedlich dotierte Gates für NMOSFET und PMOSFET einzuführen. Die Austrittsarbeitsdifferenzen von NMOSFET mit n⁺-Polysiliziumgate/p-Wanne und PMOSFET mit p⁺-Polysiliziumgate/n-Wanne sind symmetrisch zum Nullpunkt (siehe Abbildung 3.1). Mit einem p⁺-dotierten Polysiliziumgate kann also die Einsatzspannung durch eine Kanalimplantation vom n-Typ (Donatoren) angepaßt werden, und man erhält einen PMOSFET mit Oberflächenkanal. Diese Technologie wird „Dual Workfunction CMOS“ genannt. In der Literatur findet sich auch die Bezeichnung „Dual Gate CMOS“.

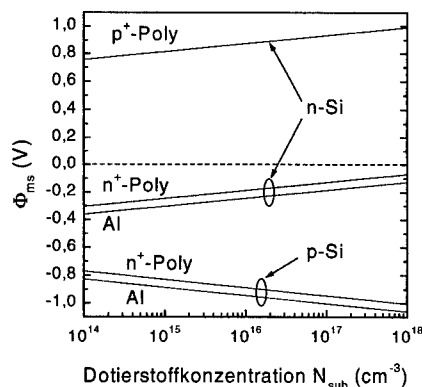


Abb. 3.1: Austrittsarbeitsdifferenz für ein n⁺-dotiertes Polysiliziumgate, ein p⁺-dotiertes Polysiliziumgate und ein Aluminiumgate in Abhängigkeit der Dotierstoffkonzentration an Donatoren (n-Si) und Akzeptoren (p-Si) im Siliziumsubstrat (Daten aus [3.2]).

3.2 Bordiffusion

Mit der Einführung eines p⁺-dotierten Polysiliziumgates für den PMOSFET ergibt sich das Problem der Bordiffusion durch das Gatedielektrikum, der sogenannten „boron penetration“. Der Segregationskoeffizient von Bor ist kleiner als eins, d.h. an der Grenzfläche zwischen Polysilizium und Siliziumdioxid kommt es im SiO₂ zu einer Anreicherung von Bor. Dieses

diffundiert durch das Gateoxid in den Kanalbereich des PMOSFETs und verursacht eine Verschiebung der Einsatzspannung zu kleineren Beträgen. In der Abbildung 3.2 ist die Abnahme der Einsatzspannung mit abnehmender Gateoxiddicke gezeigt. Die Kanalimplantation und das thermische Budget sind hier für alle Prozesse gleich. Die Substratspannung ist jeweils $V_{\text{sub}} = 0 \text{ V}$. Im Gegensatz zum NMOSFET zeigt sich für den PMOSFET eine deutliche Abweichung vom linearen Verhalten. Die Größen in der Gleichung 3.1, welche durch die Bordiffusion verkleinert werden, sind die Kanaldotierung N_{sub} , das Substratpotential Ψ_{B} und die Austrittsarbeitsdifferenz Φ_{mk} .

Die Bordiffusion in SiO_2 erfolgt vermutlich nach dem PLD-Modell [3.3]. Danach bricht Bor im Siliziumdioxid O-O-Brücken auf und bildet eine O-B-O-Bindungsstruktur. Das Bor wandert zwischen diesen Peroxidbindungsdefekten („peroxy linkage defect“, PLD) und durchdringt damit SiO_2 . Die Aktivierungsenergie für Bor in SiO_2 mit Dicken größer als 10 nm ist ca. 3,56 eV. Fluor erhöht die Bordiffusion, da es im SiO_2 zusätzliche PLDs bildet. Wasserstoff hingegen erhöht die Bordiffusion aufgrund von H-O-Si-Bindungen, die als „Hydrogen Related Defects“ (HRD) bezeichnet werden [3.3]. Bor diffundiert mit Hilfe der HRD durch die Bildung einer H-B-O-Si-Bindungsstruktur. Die Bordiffusion nimmt ab ca. 10 nm mit abnehmender Gateoxiddicke zu. Dies kann ebenfalls mit dem PLD-Modell erklärt werden. In der Übergangsschicht zwischen Silizium und SiO_2 ist die Defektdichte größer. Mit abnehmender Oxiddicke wird der Anteil der Übergangsschicht gegenüber der Oxiddicke immer größer und damit steigt die Bordiffusion an, bzw. die Aktivierungsenergie für Bordiffusion nimmt ab [3.3]. Stickstoff erhöht die Aktivierungsenergie für Bordiffusion. Im PLD-Modell wird dieser Effekt damit erklärt, daß Stickstoff, ähnlich wie Bor, Peroxidbindungsdefekte besetzt und eine O-N-O-Bindungsstruktur bildet. Damit entsteht eine Konkurrenzsituation zwischen Bor und Stickstoff.

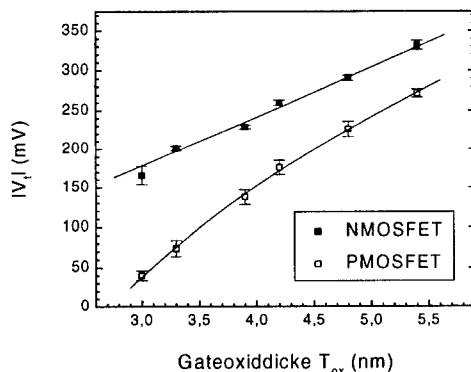


Abb. 3.2: Die Einsatzspannung des PMOSFETs nimmt nicht linear mit der Oxiddicke ab. Dies deutet auf eine zunehmende Bordiffusion hin. Die Linien dienen nur zur besseren Verdeutlichung.

Aufgrund dieser Eigenschaft von Stickstoff wurde für die CMOS-Logikgeneration mit einer minimalen Strukturbreite von $0,18\ \mu\text{m}$ eine technologische Veränderung des Gateoxides eingeführt. Durch einen Nitridierungsprozeß werden im thermisch gewachsenen SiO_2 Stickstoffatome angereichert, welche die Bordiffusion verhindern sollen.

3.3 Nitridierung von Siliziumdioxid

3.3.1 Thermische Nitridierung

Für die thermische Nitridierung von Siliziumdioxid können NH_3 (Ammoniak), N_2O (Distickstoffmonoxid, Lachgas) und NO (Stickstoffmonoxid, Stickoxid) verwendet werden. Die Pfeile in der folgenden Auflistung grundsätzlicher Nitridierungskonzepte kennzeichnen aufeinanderfolgende Einzelprozeßschritte.

- a) Thermisches $\text{SiO}_2 \rightarrow$ Temperaturschritt in NH_3 -Atmosphäre \rightarrow Freisetzen von Wasserstoff durch einen Temperaturschritt in inerter Atmosphäre
- b) Thermisches $\text{SiO}_2 \rightarrow$ Temperaturschritt in N_2O -Atmosphäre, der neben der Nitridierung auch eine weitere thermische Oxidation bewirkt \rightarrow Temperaturschritt in inerter Atmosphäre
- c) Thermische Oxidation des Siliziumsubstrates in N_2O -Atmosphäre \rightarrow Temperaturschritt in inerter Atmosphäre
- d) Thermisches $\text{SiO}_2 \rightarrow$ Temperaturschritt in reiner NO -Atmosphäre

Im Fall a) ist trotz des letzten Temperaturschrittes die Wasserstoffkonzentration im Oxid hoch. Daraus resultiert eine schlechte Oxidqualität, verursacht durch eine hohe Dichte an Haftstellen. In den Fällen b) und c) entstehen bei der Dissoziation von N_2O Sauerstoff und NO [3.2]. Der Sauerstoff führt zu einer weiteren thermischen Oxidation an der Substratoberfläche und zu einem Anstieg der Gateoxiddicke. NO ist verantwortlich für die Einbindung von Stickstoff in das Oxid und in die Übergangsschicht von Siliziumsubstrat und SiO_2 [3.4]. Der Stickstoff verhindert neben der Diffusion von Bor auch die Diffusion von Sauerstoff und hemmt die weitere thermische Oxidation des Siliziumsubstrates. Aufgrund der Erkenntnisse über die Rolle von NO wurde die Prozeßvariante d) eingeführt [3.5]. Hierbei kommt es schnell zu einer hohen Anreicherung von Stickstoff in der Grenzfläche und zu einer frühen Sättigung im Oxiddickenzuwachs bei ca. $0,4\ \text{nm}$ [3.6].

3.3.2 Implantation von Stickstoff

Eine weitere Möglichkeit der Nitridierung bietet die Implantation von Stickstoff in das Siliziumsubstrat vor der thermischen Oxidation [3.7]. Ein Nebeneffekt ist dabei die verringerte Oxidation des Substrates, da Stickstoff als Barriere gegen Sauerstoffdiffusion wirkt. Diese Variante eignet sich somit für Prozesse, die zwei verschiedene Gateoxiddicken benötigen („Dual Gate Oxide Process“). In der Abbildung 3.3 ist die Abnahme der thermisch

gewachsenen Oxiddicke für verschiedene Oxidationsprozesse zwischen 5,3 nm und 6,5 nm gezeigt. Bei hohen Dosen wird eine Sättigung erreicht. Der Effekt ist unabhängig von der Implantationsenergie, kann jedoch durch einen Temperaturschritt in inerter Atmosphäre nach der Stickstoffimplantation, also vor der thermischen Oxidation, eliminiert werden. Die Stickstoffatome werden während eines solchen Temperaturschrittes offensichtlich aus den ersten Nanometern der Substratoberfläche wieder entfernt.

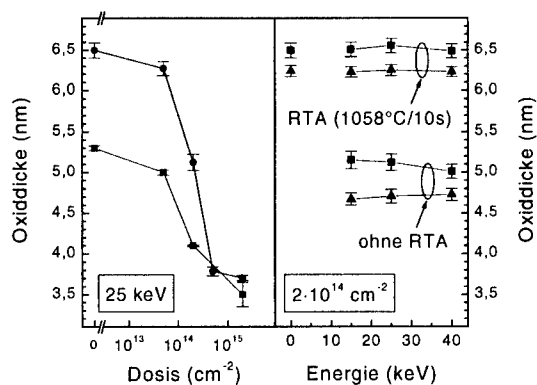


Abb. 3.3: Links: Abnahme der thermisch gewachsenen Oxiddicke (5,3 nm und 6,5 nm) durch Stickstoffimplantation in das Siliziumsubstrat vor der thermischen Oxidation. Rechts: Durch einen RTA („Rapid Thermal Anneal“) vor der Oxidation kann dieser Effekt wieder eliminiert werden.

Alle bisher beschriebenen Nitridierungsverfahren verhindern eine Verschiebung der Einsatzspannung durch eine Anreicherung von Stickstoff im Oxid und speziell in der Grenzschicht zwischen Substrat und Oxid. Eine weitere Option ist eine Diffusionsbarriere, welche bereits die Diffusion von Bor durch das Oxid verhindert. Diese Diffusionsbarriere kann z.B. durch eine Stickstoffimplantation in das Polysilizium erreicht werden [3.8, 3.9]. Die nachfolgenden Temperaturschritte im „front-end“ des weiteren Prozeßverlaufes ergeben eine Anreicherung von Stickstoff in der Grenzschicht zwischen Polysilizium und SiO₂. Bei diesem Verfahren ist darauf zu achten, daß sich keine niedrig p-dotierte Schicht im nitridierten Polysilizium ausbildet, welche eine Verarmung des Gates in Inversion zur Folge hätte („gate depletion“). In der Abbildung 3.4 sind HF-CV-Kurven gezeigt, die an diffusionsbegrenzten PMOS-Kapazitäten gemessen wurden. Die Kapazitäten im linken Teil der Abbildung haben jeweils ein stickstoffimplantiertes, p⁺-dotiertes Polysiliziumgate und das Gateoxid ist reines thermisches SiO₂ mit unterschiedlicher Dicke. Als Vergleich ist im rechten Teil der Abbildung 3.4 eine PMOS-Kapazität mit „gate depletion“ gezeigt.

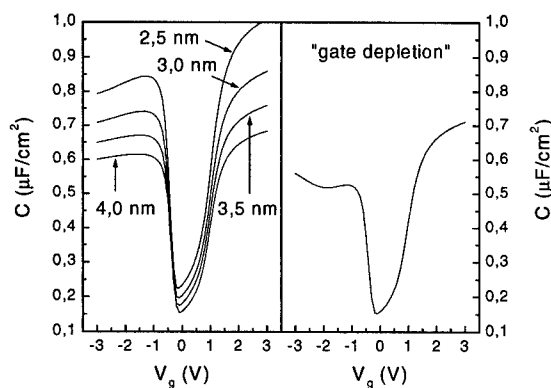


Abb. 3.4: HFCV-Messungen an diffusionsbegrenzten PMOS-Kapazitäten. Im rechten Teil sind Kapazitäten ohne „gate depletion“ und mit einem stickstoffimplantierten Polysiliziumgate gezeigt. Als Vergleich ist im rechten Teil eine PMOS-Kapazität mit „gate depletion“ aufgeführt.

3.3.3 Vergleich zwischen NO-Nitridierung und Stickstoffimplantation in das Siliziumsubstrat

Die Vorteile von NO sind eine gute Kontrollierbarkeit des Nitridierungsprozesses und der Gateoxidstärke über einen weiten Temperaturbereich. Aufgrund der stärker nitridierenden Wirkung kann ein niedrigeres thermisches Budget verwendet werden als bei N_2O . Zudem wurden für NO-nitridierte Oxide generell bessere Eigenschaften gemessen, als für vergleichbare N_2O -nitridierte Oxide [3.5]. Nachteilig ist, daß mit NO ein giftiges Gas in einer RTP-Umgebung verwendet werden muß (RTP: „Rapid Thermal Process“). Die Stickstoffimplantation in das Substrat ist aufgrund der „Dual Gate Oxide“-Prozeßvariante interessant. Es war daher das Ziel, erstmals diese beiden unterschiedlichen Prozeßvarianten direkt miteinander, sowie mit einem konventionellen Gateoxid zu vergleichen. Der Vergleich der Gateoxideigenschaften erfolgt am NMOS, um die Einflüsse von Bor im Oxid und in der Grenzfläche auszuschließen.

Das auf einem stickstoffimplantierten Substrat thermisch gewachsene Oxid sei als „NI-Oxid“ bezeichnet. Energie und Dosis der Stickstoffimplantation sind 25 keV und $2 \cdot 10^{14} \text{ cm}^{-2}$. Das thermische Oxid mit NO-Nitridierung wird nun „NO-Oxid“ genannt. Die Temperaturen und die Zeiten des NO-RTA sind $800^\circ\text{C}/900^\circ\text{C}$ und 10 s/30 s/60 s.

Die Stickstoffkonzentrationen eines NO-Oxides ($900^\circ\text{C}/30\text{s}$) und eines NI-Oxides wurden mittels SIMS-Analyse mit Sauerstoffdusche gemessen. Die Sauerstoffzufuhr gleicht die

Matrixübergänge Polysilizium/SiO₂ und SiO₂/Siliziumsubstrat aus, wie in der Abbildung 3.5 am ²⁹Si-Referenzsignal zu erkennen ist. Damit beeinflussen die Übergänge die Messung nicht. Für das 3,6 nm dicke NO-Oxid ist der Spitzenwert der Stickstoffkonzentration in der Grenzfläche zwischen Oxid und Substrat lokalisiert. Für das 3,8 nm dicke NI-Oxid liegt dieses Maximum mehr im Oxid, was durch den thermischen Oxidationsprozeß nach der Stickstoffimplantation zu erklären ist. Die Stickstoffkonzentration des NO-Oxides liegt bei 0,51 Atom-% und die des NI-Oxides bei 0,14 Atom-%. Damit ist in beiden Fällen die Anforderung einer leichten Nitridierung erfüllt. Starke Nitridierungen erhöhen nicht nur die Grenzflächenzustandsdichte, sondern verschlechtern auch die TDDB-Eigenschaften der Oxide [3.10].

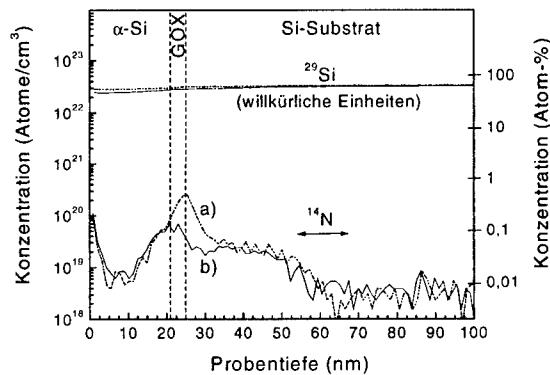


Abb. 3.5: SIMS-Analyse an nitrierten Oxiden: a) 3,6 nm dickes NO-Oxid (900°C/30s) und b) 3,8 nm dickes NI-Oxid (25 keV/2·10¹⁴ cm⁻²). Das Referenzsignal ²⁹Si dient zum Nachweis, daß die Maxima im ¹⁴N-Signal nicht durch Matrixübergänge bestimmt sind.

In der Abbildung 3.6 sind die Einsatzspannungen von NMOSFET und PMOSFET miteinander verglichen. Für das konventionelle Gateoxid zeigt sich deutlich eine Verschiebung der Einsatzspannung des PMOSFETs, verursacht durch Bordiffusion. Die leichten Nitridierungen von NO-Oxid und NI-Oxid reichen aus, um die Bordiffusion wirksam zu blockieren.

Ab einer Spitzenkonzentration von mehr als 1 Atom-% Stickstoff in der Grenzfläche ist für NO-Oxide ein starker Anstieg in der Grenzflächenzustandsdichte festzustellen [3.10]. Die Zunahme der Grenzflächenzustandsdichte während eines elektrischen Stresses wird jedoch mit zunehmender Stickstoffkonzentration kleiner und liegt unterhalb der für konventionelle

Gateoxide [3.10]. Beide Charakteristika werden auch für die vorliegenden Proben gemessen, wie der Abbildung 3.7 zu entnehmen ist.

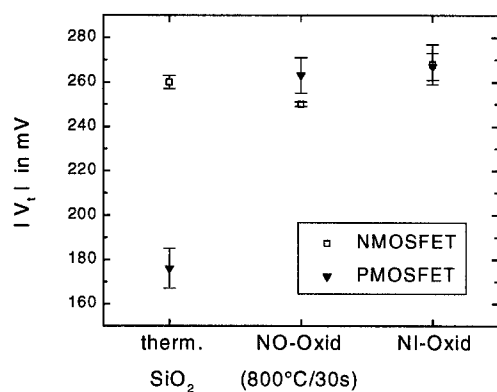


Abb. 3.6: Die nitridierten Oxide verhindern die Bordiffusion, welche für ein konventionelles Gateoxid zu einer Verschiebung der Einsatzspannung des PMOSFETs führt. Das etwas kleinere $|V_t|$ des NMOSFETs wird beim NO-Oxid vermutlich durch das abweichende thermische Budget verursacht.

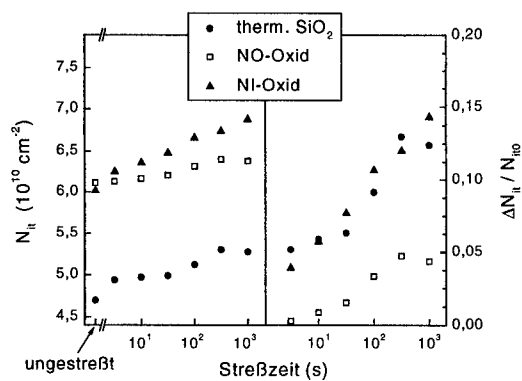


Abb. 3.7: Änderung der Grenzflächenzustandsdichte N_{it} , gemessen mit Charge-Pumping, für einen homogenen Streß an einem NMOSFET bei 27°C. Das NO-Oxid (900°C/30s) weist eine kleinere Generationsrate auf, als das konventionelle Gateoxid und das NI-Oxid ($25\text{keV}/2 \cdot 10^{14}\text{cm}^{-2}$).

Die geringere Generation von Grenzflächenzuständen kann auf die Ausbildung von Si-N-Bindungen anstelle der schwachen Si-H-Bindungen in der Grenzfläche zurückgeführt werden. Zudem verringert der Temperaturschritt in reiner NO-Atmosphäre die Wasserstoffkonzentration im Oxid, so daß auch die Ansammlung von Oxidladungen in Haftstellen kleiner ist als für ein konventionelles Oxid [3.10]. Im Gegensatz zum NO-Oxid ist die Generationsrate an Grenzflächenzuständen für das NI-Oxid in etwa gleich der des konventionellen Oxides. Eine Ursache könnte die um einen Faktor 3,6 kleinere Stickstoffkonzentration sein. Am Ende dieses Kapitels wird die Beweglichkeit der Inversionsladungsträger diskutiert und dabei eine weitere mögliche Erklärung angegeben.

Für nitridierte Oxide wird oft eine bessere Immunität gegenüber Hot-Carrier-Streß berichtet [3.5, 3.10, 3.11, 3.12]. Dies wird mit der abnehmenden Generation von Oxidladungen und Grenzflächenzuständen mit zunehmendem Stickstoffgehalt erklärt. Die vorliegenden Proben zeigen jedoch, daß nitridierte Oxide nicht in jedem Fall bessere Hot-Carrier-Eigenschaften besitzen als konventionelle Gateoxide. Die gute Vergleichbarkeit der gestreßten Transistoren ist anhand von wichtigen Parametern in der Tabelle 3.1 dokumentiert.

	Therm. Oxid	NO-Oxid (900°C/30s)	NI-Oxid (25 keV/2·10 ¹⁴ cm ⁻²)
V _t @0,1V (V)	0,296±0,004	0,248±0,001	0,300±0,001
V _t @1,8V (V)	0,267±0,004	0,219±0,002	0,271±0,003
I _d @0,1V (μA/μm)	50,2±1	46,3±0,1	49,8±0,2
I _d @1,8V (μA/μm)	339±4	334±0,7	336±0,7
Multiplikationsfaktor I _{sub} /I _d	0,021	0,019	0,022
T _{ox} (nm)	4,2	4,3	4,1

Tab. 3.1: Parameter der NMOSFETs, die für den Vergleich der Hot-Carrier-Stabilität verwendet wurden.

In den Abbildungen 3.8 und 3.9 sind die am stärksten degradierenden Größen während eines Hot-Carrier-Stresses aufgeführt. Dies sind das Maximum der Transferleitfähigkeit g_m und der lineare Drainstrom $I_{d,lin}$. Bei kurzen Streßzeiten ist die Degradation im Fall des NO-Oxides zunächst am stärksten, nähert sich bei hohen Streßzeiten jedoch der des thermischen Oxides. Die absolute Änderung der Transferleitfähigkeit Δg_m ist bei hohen Streßzeiten geringer als für das konventionelle Gateoxid. Bei hohen Streßzeiten wird also die in der Abbildung 3.7 gezeigte Eigenschaft des NO-Oxides dominant. Dabei ist zu beachten, daß beim Hot-Carrier-Streß eine wesentlich geringere Gatespannung eingestellt ist, als bei der in der Abbildung 3.7 gezeigten Messung. Der lineare Drainstrom degradiert für das NO-Oxid insgesamt am stärksten. Für das NI-Oxid ist die Degradation jeweils größer als für das thermische Oxid. Am Beispiel des NI-Oxides ist festzustellen, daß keine Verbesserung in der Hot-Carrier-Stabilität für nitridierte Oxide verbleibt, wenn die Grenzfläche nicht stabiler gegenüber der Generation von Grenzflächenzuständen wird. Gleiches gilt, wenn eine solche Stabilität insgesamt nicht

dominant ist. Das NO-Oxid zeigt, daß bei schwacher Nitridierung dieser positive Effekt erst bei hohen Streßzeiten effektiv wird, wobei eventuell bereits vorher das Degradationslimit erreicht sein kann (z.B.: $\Delta g_m/g_m = 10\%$).

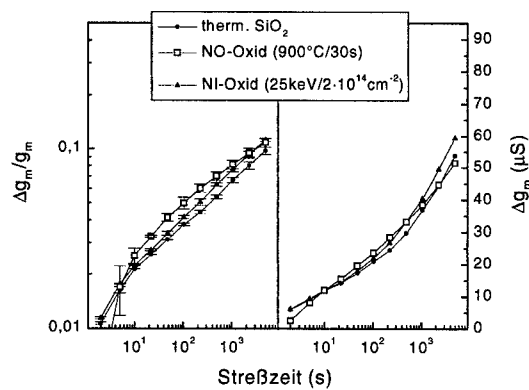


Abb. 3.8: Degradation des Maximums der Transferleitfähigkeit g_m während eines Hot-Carrier-Stresses an einem NMOSFET. Die Streßbedingungen waren: max. I_{sub} , $V_g = 1,2$ V, $V_d = 3,6$ V, $T = 27^\circ\text{C}$.

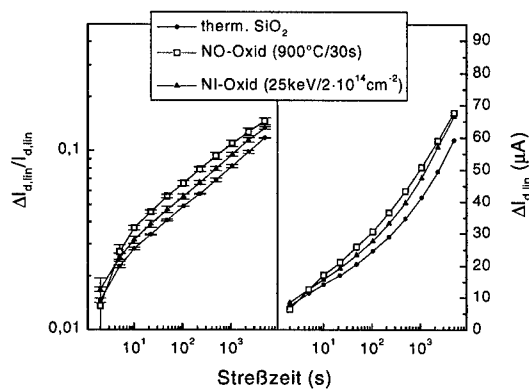


Abb. 3.9: Degradation des linearen Drainstromes $I_{d,lin}$ während eines Hot-Carrier-Stresses an einem NMOSFET. Die Streßbedingungen waren: max. I_{sub} , $V_g = 1,2$ V, $V_d = 3,6$ V, $T = 27^\circ\text{C}$.

Die Ladungsträgerbeweglichkeiten wurden anhand der Transferleitfähigkeit g_m charakterisiert. Eine Verkleinerung der Maxima der Beweglichkeiten für nitrierte Oxide ist seit Mitte der 80er Jahre bekannt [3.2]. Hori et al. entdeckten 1988 erstmals, daß in NMOS-Feldeffekttransistoren mit nitrierten Oxiden die Beweglichkeit der Elektronen bei hohen elektrischen Feldern besser ist, als für thermisches SiO_2 [3.13]. Dieses Verhalten ist besonders bei Stickstoffkonzentrationen kleiner als 2 Atom-% ausgeprägt und zeigt sich auch für die vorliegenden NO-Oxide. In der Abbildung 3.10 ist dieser flachere Verlauf von g_m bei hohen Feldern deutlich zu erkennen. Für einen qualitativen Vergleich reicht die Transferleitfähigkeit aus. Eine quantitative Analyse der Beweglichkeit müßte jedoch anhand der effektiven Beweglichkeit erfolgen (siehe Kapitel 2 und Anhang A).

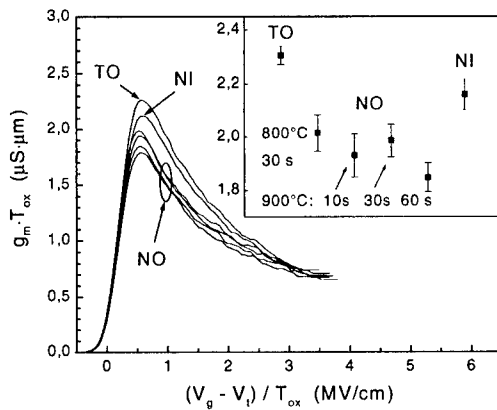


Abb. 3.10: Normierte Transferleitfähigkeit von NMOSFETs gegenüber dem vertikalen elektrischen Feld. Im Ausschnitt sind die Maximalwerte mit Standardabweichung angegeben. Die Zuordnung der Kurven erfolgt anhand dieser Maximalwerte. TO: therm. Oxid, NO: NO-Oxid, NI: NI-Oxid.

Hori erklärt diesen Effekt bei nitrierten Oxiden mit einem Grenzflächenzustandsmodell („Interface-State Model“) [3.14, 3.2]. Danach wird der unterschiedliche Einfluß der thermischen Nitridierung auf die Beweglichkeit von Elektronen und Löchern durch eine strukturelle Änderung der Grenzfläche verursacht. Diese erniedrigt die Anzahl der akzeptorartigen Haftstellen in der Grenzfläche mit Ionisationsenergien im Bereich des Leitungsbandes, nicht jedoch die Anzahl der donatorartigen Haftstellen mit Ionisationsenergien im Bereich des Valenzbandes. Für das NI-Oxid ergibt sich ein etwas höherer Spitzenwert für g_m , da der Stickstoffanteil geringer ist. Bei hohen elektrischen Feldern ist im Gegensatz zum NO-Oxid kein flacherer Verlauf zu erkennen. Damit ist auch keine ähnliche Verbesserung der Elektronenbeweglichkeit gegenüber einem thermischen Oxid zu erwarten, wie das bei thermisch nitrierten Oxiden der Fall ist. Bezogen auf das

Grenzflächenzustandsmodell von Hori ist dies ein Hinweis, daß die Grenzfläche des NI-Oxides mehr durch die thermische Oxidation nach der Stickstoffimplantation bestimmt wird. Für thermisch nitridierte Oxide findet jedoch die Nitridierung entweder nach oder während der thermischen Oxidation statt. Dabei kann sich eine signifikante Änderung der Struktur der Grenzfläche ergeben. Dies ist auch eine Erklärung dafür, daß die Generationsraten von Grenzflächenzuständen für das NI-Oxid und das thermische Oxid in etwa gleich sind (siehe Abbildung 3.7).

In der Abbildung 3.11 ist ein Vergleich zwischen PMOSFET und NMOSFET gezeigt. Für das NO-Oxid ist das unterschiedliche Verhalten der Beweglichkeiten der Inversionsladungsträger bei hohen elektrischen Feldern zu erkennen.

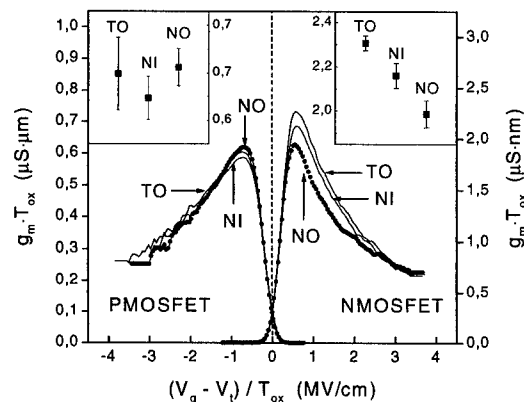


Abb. 3.11: Vergleich der normierten Transferleitfähigkeiten von NMOSFETs und PMOSFETs gegenüber dem vertikalen elektrischen Feld. Die Ausschnitte zeigen jeweils die Maximalwerte mit Standardabweichungen. TO: therm. Oxid, NO: NO-Oxid, NI: NI-Oxid.

3.3.4 Zusammenhang zwischen Bordiffusion und „Negative Bias Temperature Stress“

Durch einen „Negative Bias Temperature Stress“ (NBTS) wird die Einsatzspannung eines PMOSFETs zu mehr negativeren Werten hin verschoben. Die genaue Ursache für diese Degradation ist noch nicht geklärt. Neben der Generation von Grenzflächenzuständen spielt offensichtlich auch das Haften von Löchern im Oxid bzw. die Generation von positiven Oxidladungen eine Rolle [3.15]. Außerdem wurde gezeigt, daß Transistoren mit einem höheren thermischen Budget eine größere Verschiebung der Einsatzspannung nach NBTS

haben [3.16]. Dies wird auf eine Zunahme an positiven Oxidladungen zurückgeführt, welche durch eine höhere Borkonzentration im Oxid entsteht.

In der Abbildung 3.12 ist die NBTS-Stabilität der im vorangegangenen Kapitel analysierten Proben gezeigt. Es ist die Zeit bis zu einer Verschiebung der Einsatzspannung von $\Delta V_1 = 10$ mV gegenüber dem vertikalen elektrischen Feld aufgetragen. Ein Widerspruch scheint zu sein, daß das thermische Oxid die beste und das NO-Oxid die schlechteste Stabilität zeigt. In der Abbildung 3.13 ist die NBTS-Stabilität zweier mit N_2O nitrierter Oxide und eines thermischen Oxides gezeigt. Der Gateaufbau mit einem thermischen Oxid besitzt in diesem Fall eine Stickstoffimplantation in das Polysiliziumgate und wird hier mit „NIG-Stack“ bezeichnet. Der NIG-Stack weist die bessere NBTS-Stabilität auf.

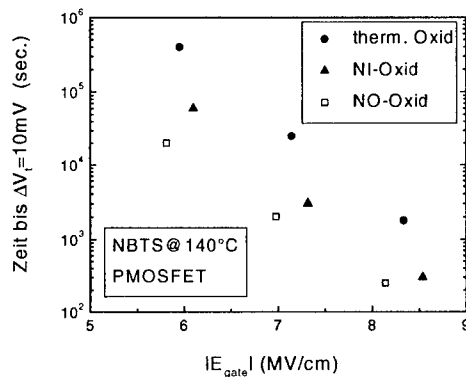


Abb. 3.12: NBTS-Stabilität von PMOSFETs mit einem NO-Oxid ($900^\circ\text{C}/30\text{s}$), einem NI-Oxid ($25\text{keV}/2 \cdot 10^{14}\text{cm}^{-2}$) und einem reinen thermischen Oxid.

Dies kann erklärt werden, wenn sowohl die Generation von Grenzflächenzuständen als auch die Generation von positiven Oxidladungen berücksichtigt wird. Befindet sich die Stickstoffbarriere gegen Bordiffusion an der Grenzfläche zwischen Substrat und Oxid, dann kommt es zu einer Anreicherung von Bor im Oxid. Bei thermisch nitrierten Oxiden wird der positive Effekt einer stabileren Grenzfläche durch die hohe Anreicherung von positiven Oxidladungen überdeckt. Im Fall eines reinen thermischen Oxides erzeugt das Bor im Oxid ebenfalls positive Oxidladungen und eine Verschiebung der Einsatzspannung zu negativeren Werten. Das Bor im Kanalbereich bewirkt allerdings eine Verschiebung der Einsatzspannung zu positiveren Werten. In diesem Fall ergibt sich also eine scheinbar bessere NBTS-Stabilität, da sich die Effekte gegenseitig kompensieren. Ist die Stickstoffbarriere gegen Bordiffusion an der Grenzfläche von Oxid und Polysilizium lokalisiert, dann erklärt sich die bessere NBTS-

Stabilität durch eine geringere Borkonzentration im Oxid. Dies ist für den NIG-Stack der Fall. Das NI-Oxid besitzt eine weniger stabile Grenzfläche als das NO-Oxid. Eine bessere NBTS-Stabilität kann somit nur erklärt werden, wenn die Bordiffusion blockiert wird, ohne explizit Bor im Oxid anzureichern. Dies wäre durch eine Nitridierung des Oxides der Fall, die von der Oxid/Poly-Grenzfläche zur Substrat/Oxid-Grenzfläche hin abnimmt. Die diskutierten Fälle sind in der Tabelle 3.2 schematisch aufgelistet.

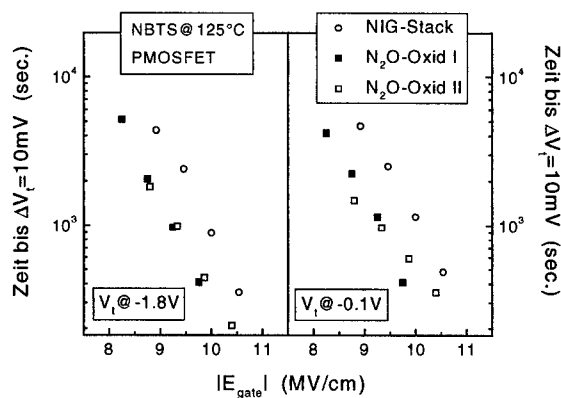


Abb. 3.13: NBTS-Stabilität von PMOSFETs mit N_2O -nitrierten Oxiden und mit einem Gateaufbau aus thermischem Oxid und einem mit Stickstoff implantierten Polysiliziumgate (NIG-Stack).

	Poly-Si	N-Profil im Oxid	Substrat	ΔV_t Bor im Kanal	ΔV_t pos. Oxid- ladungen	ΔV_t ΔD_{it}	NBTS- Stabilität
NO-Oxid N_2O -Oxid				–	negativ	sehr klein	
NIG-Stack				–	–	klein	
NI-Oxid				–	negativ klein	klein	
therm. Oxid	kein Stick- stoff			positiv	negativ	klein	

Tab. 3.2: Qualitative Erklärung für die NBTS-Stabilität von Oxiden mit unterschiedlichen Nitridierungskonzepten und von einem thermischen Oxid ohne Nitridierung.

3.4 Zusammenfassung

Mit der thermischen Nitridierung in NO-Atmosphäre und einer Stickstoffimplantation in das Siliziumsubstrat wurden zwei völlig unterschiedliche Nitridierungskonzepte miteinander verglichen. Im Gegensatz zu Darstellungen in der Fachliteratur wurde festgestellt, daß Stickstoff in der Grenzfläche zwischen SiO_2 und Substrat nicht automatisch zu einer Verbesserung der Hot-Carrier-Stabilität führt. Die bessere Stabilität der Grenzfläche, welche im allgemeinen als Grund für eine höhere Hot-Carrier-Stabilität angegeben wird, muß dafür gegenüber weiteren Degradationseinflüssen dominant werden. Im Fall des thermischen Oxides mit Stickstoffimplantation in das Substrat ergibt sich eine Stabilität gegenüber der Generation von Grenzflächenzuständen, die der eines reinen thermischen Oxides entspricht. Dies deutet auf eine gleiche Struktur der Grenzfläche hin und wird durch die Analyse der Beweglichkeit der Minoritätsladungsträger bei hohen elektrischen Feldern bestätigt.

Die Stabilität des PMOS-Feldeffekttransistors gegenüber „Negative Bias Temperature Stress“ (NBTS) ist für unterschiedliche Nitridierungskonzepte verschieden. Dies kann erklärt werden, wenn berücksichtigt wird, daß Bor im Oxid die NBTS-Stabilität verschlechtert und die Borkonzentration durch die Struktur der Stickstoffbarriere beeinflusst wird. Insgesamt ist festzustellen, daß eine Stickstoffbarriere an der Grenzfläche von Polysilizium und SiO_2 zu bevorzugen ist, da Bor in SiO_2 auch die TDDB-Eigenschaften (TDDB: „Time Dependent Dielectric Breakdown“) verschlechtert [3.17]. Dies kann durch eine Stickstoffimplantation in das Polysiliziumgate erreicht werden. Der Standardprozeß bleibt dabei bis zur Implantation des Polysiliziums gleich. Die technologische Veränderung betrifft also in diesem Fall nicht mehr den Gateoxidprozeß.

4 Das Tunneln von Ladungsträgern durch eine dünne Siliziumdioxidschicht

Bezogen auf die jeweiligen Leitungs- und Valenzbandkanten beträgt die Höhe der Potentialbarriere zwischen Siliziumdioxid und Silizium für Elektronen ca. 3,1 eV und für Löcher ca. 4,8 eV. Die Quantenmechanik erklärt eine endliche Wahrscheinlichkeit für das Tunneln von Ladungsträgern durch eine Potentialbarriere, d.h. das Oxid stellt keinen Isolator im klassischen Sinn dar. In den folgenden Kapiteln ist die Diskussion auf Elektronen beschränkt und bezieht sich ausschließlich auf eine MOS-Struktur mit entartet dotiertem Polysiliziumgate.

4.1 Tunnelstromdichte

4.1.1 Direktes Tunneln und Fowler-Nordheim-Tunneln

Direktes Tunneln liegt vor, wenn die Elektronen z.B. vom Leitungsband des Polysiliziums über die gesamte Oxiddicke in das Leitungsband des Siliziumsubstrates tunneln oder umgekehrt. Liegt über dem Oxid eine Spannung an, die größer als die Barrierenhöhe für Elektronen ist, dann tunneln die Elektronen durch eine dreieckige Potentialbarriere in das Leitungsband des Oxides. Dieser Vorgang wird als Fowler-Nordheim-Tunneln oder Fowler-Nordheim-Feldemission bezeichnet. Direktes Tunneln und Fowler-Nordheim-Tunneln sind in der Abbildung 4.1 skizziert. Mit einer Näherung nach Wentzel, Kramers und Brillouin, WKB-Näherung genannt, kann die Tunnelstromdichte durch eine beliebige Potentialbarriere analytisch bestimmt werden. Die Stromdichte für Fowler-Nordheim-Tunneln berechnet sich danach aus der bekannten Fowler-Nordheim-Gleichung [4.1].

$$\begin{aligned}
 J_{\text{FN}} &= A \cdot E_{\text{ox}}^2 \cdot \exp\left(-\frac{B}{E_{\text{ox}}}\right) \\
 A &= \frac{q^3}{16 \cdot \pi^2 \cdot \hbar \cdot \Phi_b} \\
 B &= \frac{4 \cdot \sqrt{2 \cdot m_{\text{ox}}}}{3 \cdot \hbar \cdot q} \cdot \Phi_b^{3/2}
 \end{aligned} \tag{4.1}$$

Alle Größen sind in SI-Einheiten angegeben, die Potentialbarriere Φ_b also in Joule. E_{ox} ist das elektrische Feld im Oxid und m_{ox} die effektive Elektronenmasse in der Bandlücke von SiO_2 . Die analytische Gleichung für den direkten Tunnelstrom findet sich z.B. in [4.2, 4.3].

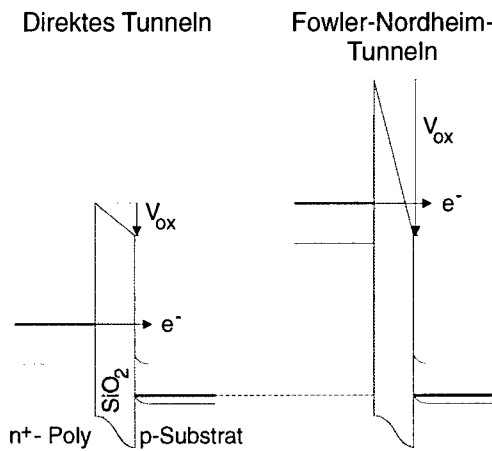


Abb. 4.1: Direktes Tunneln und Fowler-Nordheim-Tunneln von Elektronen am Beispiel einer NMOS-Struktur mit einem n^+ -dotierten Polysiliziumgate. Es ist jeweils der ideale Potentialverlauf im Oxid skizziert.

4.1.2 Berechnung mit einem Transmissionsmatrixansatz

Eine beliebige Potentialbarriere kann durch ein Stufenpotential angenähert werden. Analog zum Tunneln durch eine rechteckige Potentialbarriere (siehe z.B. [4.4]) ist die Wellenfunktion eines Elektrons mit der Energie E in der j -ten Potentialstufe

$$\Psi_j(x) = A_j \cdot \exp(ik_j x) + B_j \cdot \exp(-ik_j x) \quad (4.2)$$

$$k_j = \frac{1}{\hbar} \cdot \sqrt{2 \cdot m_j \cdot (E - U_j)} \quad (4.3)$$

U_j ist die Potentialhöhe der j -ten Stufe und m_j ist die effektive Masse des Elektrons in dieser Stufe. Die Randbedingungen für die Stetigkeit dieser stationären Lösungen der Schrödingergleichung und deren Ableitungen ergibt ein Gleichungssystem zur Bestimmung der Faktoren A_j und B_j . Nach Ando et al. [4.5] kann dieses Gleichungssystem auf eine Multiplikation von 2×2 -Matrizen zurückgeführt werden. Aus dem Produkt aller 2×2 -Matrizen läßt sich die Transmissionsamplitude berechnen. Damit ist es möglich, sowohl die Transmissionswahrscheinlichkeit als auch die Stromdichte für Elektronen zu bestimmen. Die Rechnung ist korrekt [4.6] und berücksichtigt im Gegensatz zur WKB-Näherung auch alle Reflexionen, die zur Beschreibung von quantenmechanischen Resonanz- und Interferenzphänomenen notwendig sind. Zudem ist der Algorithmus von Ando et al. für eine einfache Berechnung der Tunnelstromdichte geeignet [4.7]. In den folgenden Kapiteln wird u.a. gezeigt, daß es für diesen Algorithmus einen konsistenten Parametersatz gibt, der es ermöglicht, experimentelle Tunnelstromdaten bei hohen Spannungen sehr gut zu simulieren.

4.2 Gateleckströme in „Dual Workfunction CMOS“-Technologien

Es wird im weiteren von einer diffusionsbegrenzten Kapazität oder einem Transistor ausgegangen. Das Substrat, bzw. der Wannenkontakt, und die Diffusionskontakte liegen auf Masse. Für Dual-Workfunction-CMOS müssen vier Fälle unterschieden werden:

- 1) NMOS mit negativer Gatespannung, d.h. Elektroneninjektion in das Oxid vom n⁺-dotierten Polysiliziumgate aus („Gate Injection“, GI).
- 2) NMOS mit positiver Gatespannung, d.h. Elektroneninjektion in das Oxid vom p-Substrat aus („Substrate Injection“, SI).
- 3) PMOS mit negativer Gatespannung, d.h. Elektroneninjektion in das Oxid vom p⁺-dotierten Polysiliziumgate aus („Gate Injection“, GI).
- 4) PMOS mit positiver Gatespannung, d.h. Elektroneninjektion in das Oxid vom n-Substrat aus („Substrate Injection“, SI).

4.2.1 Substrat in Akkumulation

Im Substrat liegt in den Fällen 1) und 4), also für NMOS/GI und PMOS/SI, Akkumulation vor. Für den Betrag der Gatespannung V_g gilt jeweils:

$$|V_g| = |V_{ox}| + |\Phi_{ms}| + |V_{Si}| + \Delta\Phi_b = |V_{ox}| + \frac{E_{gap}}{q} + \Delta\Phi_b \quad (4.4)$$

V_{ox} ist die Oxidspannung, Φ_{ms} ist die Austrittsarbeitsdifferenz zwischen Gate und Substrat, V_{Si} ist die Bandverbiegung im Substrat, $\Delta\Phi_b$ ist der Unterschied zwischen den Barrierenhöhen für Polysilizium/SiO₂ und für Siliziumsubstrat/SiO₂ und E_{gap} ist die Bandlücke von Silizium. Dabei wird angenommen, daß die Kanten des jeweiligen Bandes der Majoritätsladungsträger in starker Akkumulation am Fermi-niveau fixiert sind, so daß gilt:

$$|\Phi_{ms}| + |V_{Si}| = \frac{E_{gap}}{q} \quad (4.5)$$

Dies ist in der Abbildung 4.2 für die Fälle „Akkumulation“ dargestellt. Die Berechnung der Tunnelstromkurven erfolgt in Abhängigkeit von der Oxidspannung V_{ox} . Um die gemessenen Tunnelstromkurven in derselben Graphik darzustellen, muß also vom Betrag der Gatespannung ein Betrag ΔV subtrahiert werden.

Mit Hilfe der Oszillationen im Tunnelstrom kann eine sehr gute Anpassung der Parameter für die Berechnung der Tunnelstromdichten erreicht werden. Die Oszillationen entstehen, wenn nach dem Fowler-Nordheim-Tunneln der Elektronentransport im Leitungsband des Oxides ballistisch ist, also kein inelastischer Stoß auftritt. Die Wellenfunktion wird am Potentialsprung zwischen den Leitungsbandkanten von Oxid und Anode reflektiert und kann mit der einfallenden Welle des ballistischen Elektrons interferieren [4.8]. Die Distanz, welche

das Elektron im Oxidleitungsband zurücklegen muß, ist abhängig von der Oxidspannung. Bei bestimmten Spannungswerten kommt es somit zu einer positiven Interferenz, also zu einem lokalen Maximum im Tunnelstrom und umgekehrt. Dieses quantenmechanische Phänomen wird für Oxide im Bereich von ca. 3-6 nm gemessen. Es muß dafür die mittlere freie Weglänge zwischen zwei inelastischen Stößen für Elektronen im Leitungsband von SiO_2 unterschritten werden. Zudem muß am Oxid noch eine Spannung größer als 3,1 V angelegt werden können, um Fowler-Nordheim-Tunneln zu erreichen.

Es wurde ein konsistenter Parametersatz gefunden, der bei hohen Spannungen eine sehr gute Übereinstimmung mit den Meßdaten für alle vier Fälle und für verschiedene Oxiddicken liefert. Dieser Parametersatz ist in der Tabelle 4.1 aufgelistet. Für die effektive Elektronenmasse in der Bandlücke von SiO_2 werden in der Literatur Werte zwischen $0,3 \cdot m_e$ und $0,5 \cdot m_e$ angegeben [4.9]. Eine effektive Masse von $0,5 \cdot m_e$ ergibt meist die beste Übereinstimmung für Fowler-Nordheim-Tunnelstromdaten, ebenso wie eine Barrierenhöhe zwischen dem Siliziumsubstrat und SiO_2 von 3,1 eV. Für den Fall, daß das $\langle 100 \rangle$ -orientierte Siliziumsubstrat die Kathode bildet, wurde für die effektive Masse in der Kathode der longitudinale Wert 0,98 gewählt. Vereinfachend wurde für die Polysiliziumkathode der Mittelwert aus den beiden transversalen und der longitudinalen Masse $(0,98 + 2 \cdot 0,19)/3 = 0,45$ verwendet. Die effektive Masse im Leitungsband von SiO_2 wurde angepaßt und ist mit 0,55 etwas kleiner als andere veröffentlichte Werte [4.10]. Beim Übergang in das Leitungsband der Anode ergaben sich für effektive Massen deutlich kleiner als 1 hohe Amplituden der Oszillationen, die mit den Meßwerten nicht vereinbar waren. Für die Barrierenhöhe zwischen Polysilizium und SiO_2 finden sich in der Literatur Werte zwischen 2,9 eV und 3,1 eV. Die besten Anpassungen wurden hier für 3,02 eV erreicht. Für Messungen und Rechnungen gilt immer: $T = 300 \text{ K}$.

	NMOS/GI	NMOS/SI	PMOS/GI	PMOS/SI
Zustand im Substrat	Akkumulation	Inversion	Inversion	Akkumulation
Barrierenhöhe Φ_b (eV) für Poly-Si/ SiO_2	3,02	—	3,02	—
Barrierenhöhe Φ_b (eV) für Si/ SiO_2	—	3,10	—	3,10
Effektive Masse in der Kathode	0,45	0,98	0,45	0,98
Effektive Masse in der Bandlücke von SiO_2	0,5	0,5	0,5	0,5
Effektive Masse im Leitungsband von SiO_2	0,55	0,55	0,55	0,55
Effektive Masse im Leitungsband der Anode	1	1	1	1
$\Delta V = V_g - V_{ox} $ (V)	1,15	—	—	0,97
$\Delta V = V_g - V_{ox} - V_{poly} $ (V)	—	0,23	0,21	—
Dotierung des Polysiliziumgates N_{poly} (cm^{-3})	—	$1 \cdot 10^{20}$	$8 \cdot 10^{19}$	—
Bemerkung			„Interface State Injection Model“	

Tab. 4.1: Konsistenter Parametersatz zur Berechnung der Tunnelstromdichte. Die effektiven Massen des Elektrons sind als Vielfache der freien Elektronenmasse m_e angegeben.

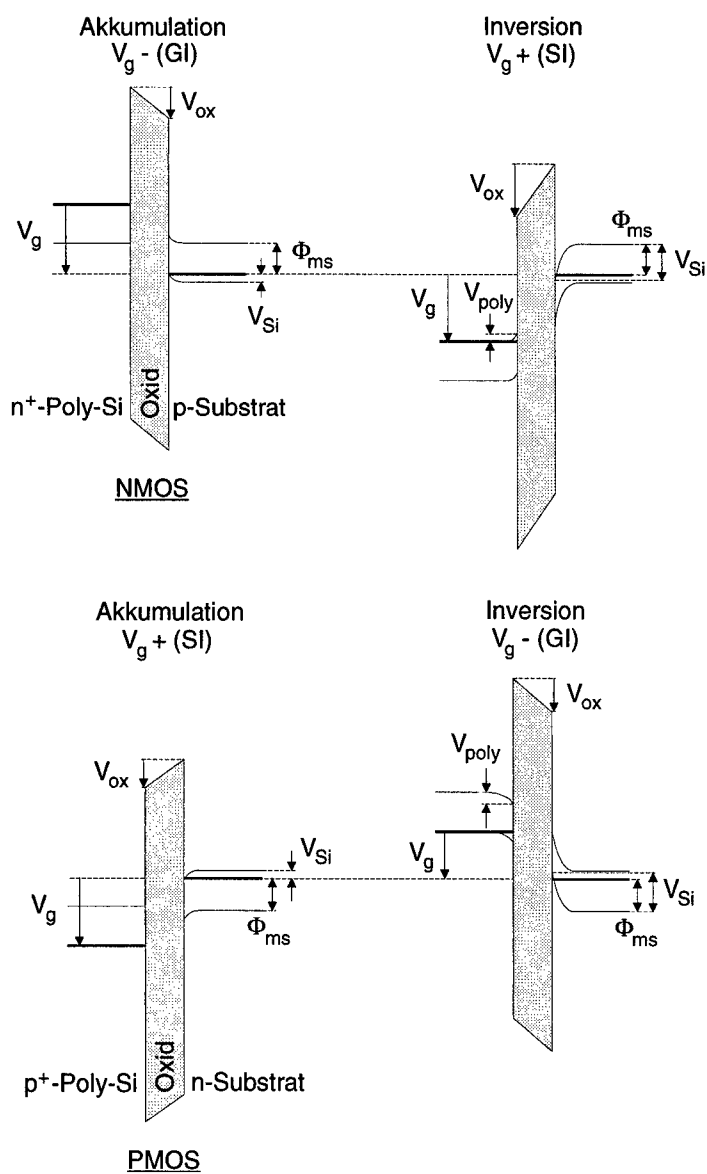


Abb. 4.2: Energiediagramme für die vier Fälle, die bei Dual-Workfunction-CMOS unterschieden werden müssen. Das Fermi-niveau ist durch eine dickere Linie gekennzeichnet. Die Beschriftung bezieht sich jeweils auf die Potentialdifferenzen.

In den Abbildungen 4.3 und 4.4 sind die gemessenen und die berechneten Stromdichten für drei verschiedene Oxiddicken dargestellt. Der ermittelte Betrag von ΔV ist für NMOS/GI gleich 1,15 V und für PMOS/SI gleich 0,97 V. Mit $E_{\text{gap}} = 1,12$ eV wäre demnach die Differenz der Barrierenhöhen $\Delta\Phi_b$ für den Fall NMOS/GI gleich 0,03 V und für den Fall PMOS/SI gleich -0,05 V. Die Abweichung von den Werten für die verwendeten Barrierenhöhen ist im Rahmen der experimentellen Genauigkeit.

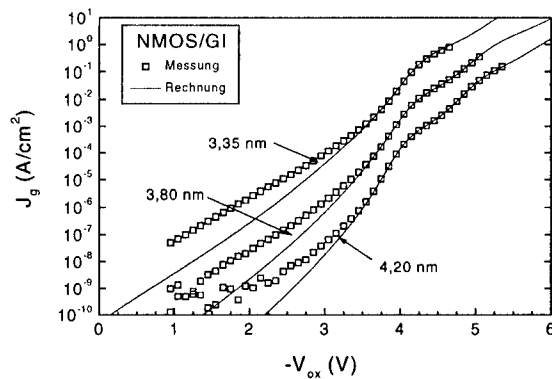


Abb. 4.3: Gemessene und berechnete Tunnelstromdichten für den Fall NMOS/GI.

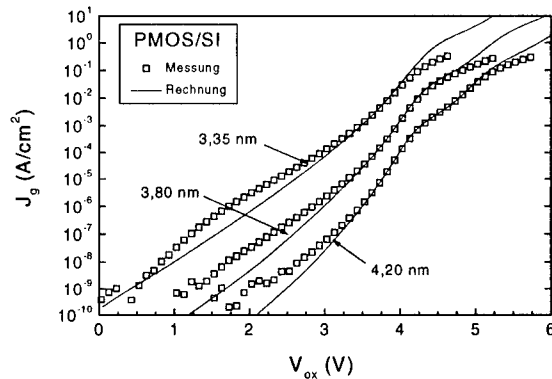


Abb. 4.4: Gemessene und berechnete Tunnelstromdichten für den Fall PMOS/SI. Bei ca. $0,1 \text{ A/cm}^2$ tritt hier ein Serienwiderstand auf, der den gemessenen Strom gegenüber der exakten Rechnung verkleinert.

Die Abweichung der Meßdaten von der Rechnung bei kleinen Spannungen wird durch einen erhöhten Transport von Elektronen mittels Ladungen oder Haftstellen im Oxid verursacht. Dieser Transport wird in der Literatur als „Trap-Assisted Tunneling“ (TAT) bezeichnet und ergibt eine höhere Stromdichte im Vergleich zum reinen direkten Tunneln. „Trap-Assisted Tunneling“ wird im Kapitel 4.4 genauer diskutiert. In allen Rechnungen wird der Schottky-Effekt, also die Verkleinerung der Potentialbarriere durch die attraktive Kraft von Spiegelladungen in der Kathode, nicht berücksichtigt. Diese Verkleinerung der Potentialbarriere nimmt proportional zur Quadratwurzel des elektrischen Feldes im Oxid zu [4.11]. Damit würde die Tunnelstromdichte mit zunehmendem elektrischen Feld stärker ansteigen, als dies für eine konstante Barrierenhöhe der Fall ist. Aus der guten Übereinstimmung zwischen Rechnung und Messung ist daher zu schließen, daß die durch den Schottky-Effekt verursachte Abhängigkeit der Barrierenhöhe vom elektrischen Feld vernachlässigt werden kann.

4.2.2 Substrat in Inversion

In den Fällen 2) und 4) ist das Substrat invertiert. Dies sind die Betriebsbedingungen für den jeweiligen MOSFET, also NMOS/SI und PMOS/GI. Im Gegensatz zum Akkumulationsfall tritt hier eine Bandverbiegung im Polysilizium auf, die als V_{poly} bezeichnet wird. Für den Betrag der Gatespannung gilt nun:

$$|V_g| = |V_{\text{ox}}| - |\Phi_{\text{ms}}| + |V_{\text{Si}}| + |V_{\text{poly}}| + \Delta\Phi_b \quad (4.6)$$

Die Bandverbiegung im Polysilizium ist abhängig von der Oxidspannung V_{ox} . Es gilt die bekannte Gleichung für die Bandverbiegung an einer Halbleiteroberfläche [4.12]:

$$V_{\text{poly}} = \frac{\epsilon_{\text{ox}}^2 \cdot E_{\text{ox}}^2}{2 \cdot q \cdot \epsilon_{\text{Si}} \cdot N_{\text{poly}}} = \frac{\epsilon_{\text{ox}}^2 \cdot V_{\text{ox}}^2}{2 \cdot q \cdot \epsilon_{\text{Si}} \cdot N_{\text{poly}} \cdot T_{\text{ox}}} = K \cdot V_{\text{ox}}^2 \quad (4.7)$$

Dies muß bei der Berechnung der Tunnelstromdichte berücksichtigt werden. Der Zusammenhang zwischen der Gatespannung V_g und der über dem Oxid abfallenden Spannung V_{ox} ergibt folgende quadratische Gleichung:

$$K \cdot V_{\text{ox}}^2 + |V_{\text{ox}}| - |V_g| = 0 \quad (4.8)$$

Damit gilt:

$$V_{\text{ox}}(V_g) = \frac{-1 + \sqrt{1 + 4 \cdot K \cdot |V_g|}}{2 \cdot K} \quad (4.9)$$

Angenommen, die Kante des jeweiligen Bandes der Minoritätsladungsträger ist in starker Inversion am Fermi-niveau fixiert, dann gilt:

$$-|\Phi_{\text{ms}}| + |V_{\text{Si}}| = 0 \quad (4.10)$$

Der einzige verbleibende konstante Spannungsbetrag wäre somit die Differenz der Barrierenhöhen $\Delta\Phi_b$, wobei diese nun für den Fall NMOS/SI negativ und für den Fall

PMOS/GI positiv ist. Es zeigt sich jedoch, daß die Gatespannung der experimentellen Daten um einen konstanten Betrag korrigiert werden muß, der größer als $\Delta\Phi_b$ ist. Dieser Betrag ist in der Tabelle 4.1 unter ΔV aufgelistet. Der höhere Betrag ist darauf zurückzuführen, daß die Kante des jeweiligen Bandes der Minoritätsladungsträger in starker Inversion nicht am Fermi-niveau fixiert ist. Die Bandverbiegung im Substrat V_{Si} ist somit größer als die Differenz der Austrittsarbeiten Φ_{nk} . In der Abbildung 4.2 sind die Energiediagramme für die Fälle „Inversion“ skizziert.

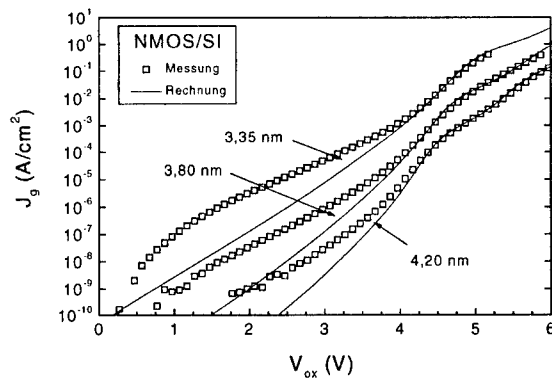


Abb. 4.5: Gemessene und berechnete Tunnelstromdichten für den Fall NMOS/SI.

Aufgrund der Gleichung 4.9 ergibt sich eine andere Abhängigkeit der Stromdichte von der Gatespannung als bei den Akkumulationsfällen. Die Bandverbiegung im Polysilizium ist dabei signifikant für den ansteigenden Verlauf der Tunnelstromdichte mit der Gatespannung. Somit läßt sich die Dotierung im Polysilizium N_{poly} sehr gut über die Anpassung der berechneten Daten bestimmen. Der ermittelte Wert ist in der Tabelle 4.1 aufgeführt. In der Abbildung 4.5 sind die gemessenen und die berechneten Stromdichten für den Fall NMOS/SI gezeigt.

In der Abbildung 4.6 ist eine Zusammenstellung der gemessenen Stromdichten für alle vier Fälle gezeigt. Im Gegensatz zu den Akkumulationsfällen NMOS/GI und PMOS/SI sind die Inversionsfälle NMOS/SI und PMOS/GI nicht umgekehrt äquivalent zueinander. Mit den obigen Gleichungen und dem bisherigen Algorithmus kann der Fall PMOS/GI noch nicht ausreichend erklärt werden. Dies wird im Kapitel 4.3.2 anhand des Vergleiches von Rechnungen und Messung belegt.

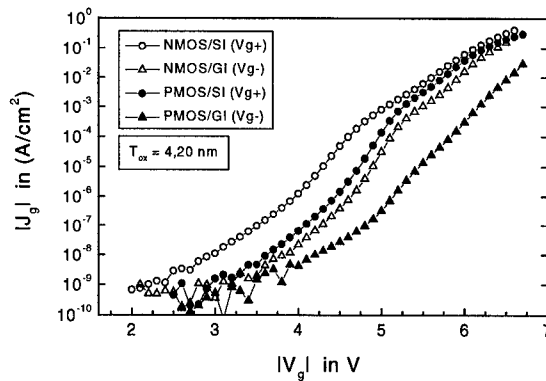


Abb. 4.6: Experimentell gemessene Stromdichten für ein 4,20 nm dickes Oxid. Im Fall NMOS/SI zeigt sich der flachere Verlauf, welcher durch die Bandverbiegung im Polysilizium verursacht wird. Die $J_g(V_g)$ -Charakteristik im Fall PMOS/GI unterscheidet sich deutlich von der im Fall NMOS/SI.

4.3 Das Tunneln von Elektronen am Beispiel des PMOS und negative Gatespannungen

Das Polysiliziumgate des PMOSFETs ist p^+ -dotiert. Die Dichte an Elektronen im Leitungsband des Gates ist daher extrem klein. Es muß erwartet werden, daß im Fall PMOS/GI Elektronen auch vom Valenzband aus in das Oxid injiziert werden. Im folgenden wird ein neues Modell eingeführt [4.13], das den Gatestrom im Fall PMOS/GI beschreiben kann.

4.3.1 „Interface State Injection Model“

Für eine Elektroneninjektion von der Valenzbandkante des Polysiliziums in das Oxid beträgt die Höhe der Potentialbarriere $\Phi_b + 1,12 \text{ eV} \approx 4,1 \text{ eV}$. Eine einfache Berechnung mit dieser Barrierenhöhe kann die gemessene Stromdichte nicht beschreiben. Das neu entwickelte „Interface State Injection Model“ [4.13] bezieht sich auf die obere Grenze für das Energieniveau in der Kathode, von dem aus die Elektronen tunneln können. Es wird angenommen, daß weder die Valenzbandkante noch die Leitungsbandkante, sondern das Fermi-niveau dafür maßgebend ist. In der Bandlücke existieren Energieniveaus von Grenzflächenzuständen, die bis zum Fermi-niveau mit Elektronen besetzt sind. Somit stellen sie eine geeignete Quelle für die Injektion von Elektronen dar. Die Besetzung der Grenzflächenzustände erfolgt über quantenmechanisches Tunneln von Elektronen aus dem

Valenzband. Dies ist analog zum Band-Band-Tunneln im Fall von „Gate-Induced Drain Leakage“ (GIDL, siehe Anhang A). Die Barrierenhöhe hierfür ist die Bandlücke von Silizium E_{gap} und die maximale Tunneldistanz ist durch die maximale Weite der Verarmungszone gegeben. Aufgrund der Bandkrümmung durchtunnelt jedoch der Großteil der Elektronen eine kleinere Distanz. Für die maximale Weite der Verarmungszone w_m gilt [4.2]:

$$w_m = \sqrt{\frac{2 \cdot \epsilon_{\text{Si}} \cdot 2 \cdot \Psi_B}{q \cdot N_{\text{dop}}}} \quad (4.11)$$

$$2 \cdot \Psi_B = \frac{2 \cdot kT}{q} \ln \left(\frac{N_{\text{dop}}}{n_i} \right) \quad (4.12)$$

N_{dop} ist die Dotierstoffkonzentration, und n_i ist die intrinsische Ladungsträgerkonzentration. Im Fall des entartet dotierten Polysiliziumgates beträgt die maximale Weite der Verarmungszone wenige Nanometer: für $N_{\text{dop}} = 8 \cdot 10^{19} \text{ cm}^{-3}$ ist $w_m = 4,4 \text{ nm}$. In der Abbildung 4.7 wird die Tunnelstromdichte mittels der Fowler-Nordheim-Tunnelstromdichte (siehe Gleichung 4.1) abgeschätzt. Es sind die Stromdichten für die Tunneldistanzen 4,4 nm und 1,0 nm gezeigt. Im linken Teil der Abbildung 4.7 ist die Näherung der Potentialbarriere schematisch skizziert. Die Potentialdifferenz für die Fowler-Nordheim-Gleichung ändert sich bei dieser Abschätzung mit der Tunneldistanz. Aus den jeweiligen Stromdichten ist zu schließen, daß bei hohen Dotierungen eine permanente Besetzung der Grenzflächenzustände durch das Tunneln von Elektronen gegeben ist.

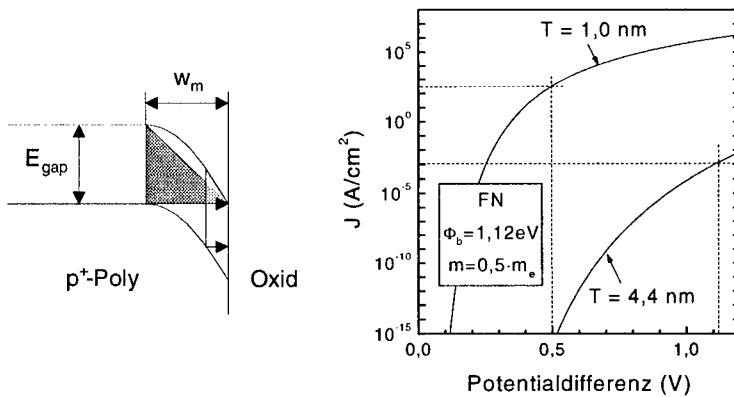


Abb. 4.7: Abschätzung der Tunnelstromdichte an Elektronen vom Valenzband des entartet dotierten Polysiliziumgates (PMOS/GI) in die Grenzflächenzustände. Als Abschätzung dient die Fowler-Nordheim-Tunnelstromdichte (FN). Die maximale Tunneldistanz ist 4,4 nm und die Barrierenhöhe entspricht der Bandlücke E_{gap} .

Demgegenüber ist die Zeitkonstante τ_{it} der thermischen Besetzung der Grenzflächenzustände zu groß, um eine ausreichende Stromdichte durch das Oxid erklären zu können. Die Zeitkonstante beträgt an der Bandkante ca. 10 μ s und in der Mitte der Bandlücke ca. 10 ms [4.2]. Für die maximal mögliche Stromdichte gilt:

$$J = \frac{q \cdot N_{it}}{\tau_{it}} \quad (4.13)$$

Damit liegt bei einer Grenzflächenzustandsdichte von $N_{it} = 10^{10} \text{ cm}^{-2}$ die maximal mögliche Stromdichte gemäß Gleichung 4.13 zwischen $1,6 \cdot 10^{-7} \text{ A/cm}^2$ und $1,6 \cdot 10^{-4} \text{ A/cm}^2$.

Das Fermi-niveau bestimmt also das Energieniveau und damit die Barrierenhöhe für die Injektion von Elektronen in das Oxid. Die Höhe der Potentialbarriere wird daher im neu entwickelten „Interface State Injection Model“ spannungsabhängig definiert und berechnet sich für das p⁺-dotierte Polysiliziumgate aus der folgenden Gleichung:

$$\Phi(V) = \Phi_b + E_{\text{gap}} - q \cdot V_{\text{poly}}(V) \quad (4.14)$$

Dabei ist Φ_b die reguläre Barrierenhöhe, also die Differenz zwischen den Leitungsbandkanten von Polysilizium und SiO₂. E_{gap} ist die Bandlücke von Silizium und V_{poly} die Bandverbiegung im Polysiliziumgate. In der Abbildung 4.6 ist für den Fall PMOS/GI zu erkennen, daß der $J_g(V_g)$ -Kurvenverlauf deutlich steiler ansteigt als für NMOS/SI. Es wird in den folgenden Kapiteln gezeigt, daß der steilere Kurvenverlauf mit der spannungsabhängigen Barrierenhöhe des „Interface State Injection Model“ erklärbar ist.

4.3.2 Berechnung mit dem Transmissionsmatrixansatz

Die spannungsabhängige Barrierenhöhe wurde in den Algorithmus zur Berechnung der Tunnelstromdichte implementiert. Die Modifikation ergänzt die im vorangegangenen Kapitel beschriebenen Änderungen für den Inversionsfall. In der Abbildung 4.8 ist die Berechnung für das 4,2 nm dicke Gateoxid mit einfachen Berechnungen ohne Berücksichtigung des „Interface State Injection Model“ verglichen. Letztere beinhalten nur die Modifikationen für den Inversionsfall und eine konstante Barrierenhöhe von 3,02 eV, die der Leitungsbandkante entspricht, bzw. 4,14 eV, die der Valenzbandkante entspricht. Die Abbildung 4.8 bestätigt zum einen, daß die einfache Rechnung den Fall PMOS/GI nicht erklären kann. Zum anderen zeigt sie die erfolgreiche Beschreibung mit Hilfe des „Interface State Injection Model“. Es ist auch zu erkennen, daß die Messung tatsächlich zwischen den beiden Grenzfällen der Elektroneninjektion verläuft, also der Injektion ausgehend von der Valenzbandkante und der Injektion ausgehend von der Leitungsbandkante.

In der Abbildung 4.9 sind die gemessenen und berechneten Stromdichten für die drei verschiedenen Oxiddicken gezeigt. Die Übereinstimmung bei verschiedenen Oxiddicken und unter Verwendung des konsistenten Parametersatzes bestätigt erneut das eingeführte „Interface State Injection Model“. Der Unterschied zwischen den Gateleckströmen für die Fälle NMOS/SI und PMOS/GI (vgl. Abbildung 4.6) wird also nicht durch eine

Spannungsverschiebung verursacht, wie u.a. in der Literatur angenommen wird [4.14]. Aufgrund der höheren Potentialbarriere verringert sich die Tunnelstromdichte.

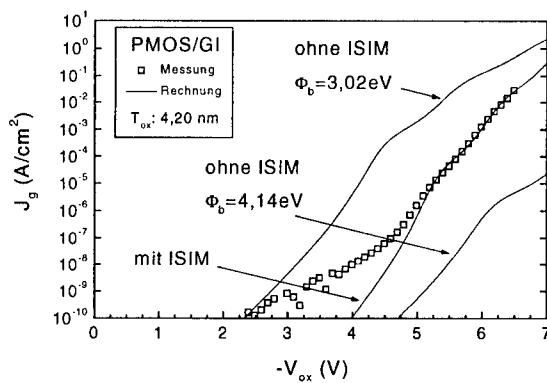


Abb. 4.8: Vergleich zwischen Messung und Rechnungen mit und ohne Berücksichtigung des „Interface State Injection Model“ (ISIM).

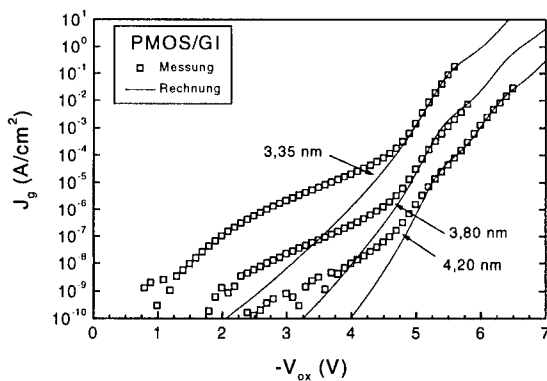


Abb. 4.9: Gemessene und berechnete Tunnelstromdichten für den Fall PMOS/GI.

4.3.3 Berechnung mit der Fowler-Nordheim-Gleichung

Das „Interface State Injection Model“ kann auch sehr einfach in eine Berechnung mit der Fowler-Nordheim-Gleichung implementiert werden. Die Faktoren A und B in der Gleichung 4.1 werden spannungsabhängig, da beide aus der spannungsabhängigen Barrierenhöhe berechnet werden. Die Rechnung und die Messung sind anhand des 4,20 nm dicken Oxides in der Abbildung 4.10 miteinander verglichen. Die in der Abbildung 4.10 gewählte Darstellung wird als Fowler-Nordheim-Graphik bezeichnet. Für eine konstante Höhe der Potentialbarriere stellt die Fowler-Nordheim-Tunnelstromdichte darin eine Gerade mit der Steigung B dar. Die Berechnung mit dem „Interface State Injection Model“ (ISIM) folgt sehr gut dem Verlauf der Meßdaten. Die Meßdaten weisen auch in der Fowler-Nordheim-Graphik die typischen Oszillationen auf. Eine an die Rechnung angepaßte Gerade verdeutlicht, daß die Rechnung zu keiner Geraden in der Fowler-Nordheim-Graphik führt, da ja die Höhe der Barriere im „Interface State Injection Model“ spannungsabhängig ist.

In der Abbildung 4.10 sind auch Berechnungen ohne dem „Interface State Injection Model“ (ISIM) gezeigt. Diese Rechnungen berücksichtigen wieder die Barrierenhöhen, welche den beiden Bandkanten entsprechen. Aus der jeweiligen Lage und den Steigungen dieser Rechnungen ist zu entnehmen, daß sich mit zunehmendem elektrischen Feld die Injektion von Elektronen aus der Nähe der Valenzbandkante in Richtung der Leitungsbandkante verschiebt. Dies konnte bereits aus der Abbildung 4.8 geschlossen werden und bestätigt das „Interface State Injection Model“.

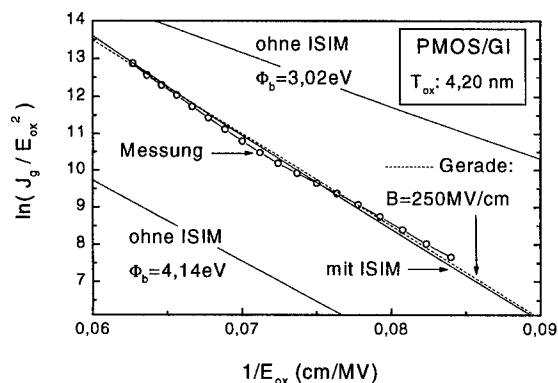


Abb. 4.10: Fowler-Nordheim-Graphik mit den Meßdaten des 4,20 nm dicken Gateoxides und den Fowler-Nordheim-Rechnungen mit und ohne „Interface State Injection Model“ (ISIM). Die Anpassung einer Geraden an die ISIM-Rechnung ergibt ein mittleres B von 250 MV/cm.

4.4 Weitere Gateleckstrommechanismen am Beispiel des PMOS und negative Gatespannungen

4.4.1 „Trap-Assisted Tunneling“

Vergleicht man die Stromdichten bei kleinen Spannungen für die vier Fälle, die in den Abbildungen 4.3, 4.4, 4.5 und 4.9 gezeigt sind, dann fällt auf, daß die Differenz zwischen Rechnung und Messung groß ist, insbesondere für den Fall PMOS/GI. Für die kleinste Oxiddicke von 3,35 nm ist dies sehr gut zu erkennen. In der Literatur herrscht Einigkeit darüber, daß die Ursache für den erhöhten Gateleckstrom Haftstellen im Oxid sind [4.15, 4.16]. Der Leckstrommechanismus wird allgemein als „Trap-Assisted Tunneling“ (TAT) bezeichnet. Dabei ist zwischen dem ungestreßten Oxid und dem Strom nach elektrischem Streß, welcher als „Stress-Induced Leakage Current“ (SILC) bezeichnet wird, zu unterscheiden. Im folgenden wird auf das Thema „Trap-Assisted Tunneling“ eingegangen, während sich das nächste Kapitel und das Kapitel 5.1 mit dem SILC befassen.

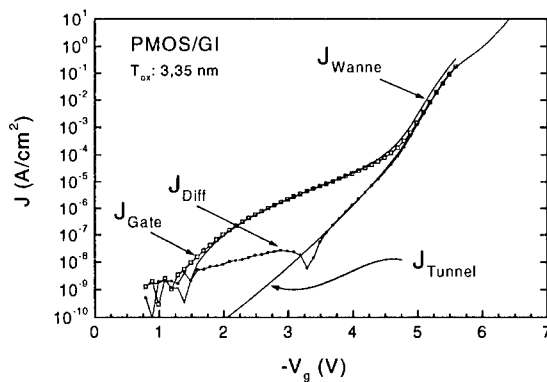


Abb. 4.11: Stromdichten gegenüber der am Gate angelegten Spannung im Fall PMOS/GI. J_{Wanne} , J_{Diff} und J_{Gate} sind die am Wannenkontakt, am Diffusionskontakt und am Gate gemessenen Stromdichten. J_{Tunnel} ist die im Kapitel 4.3.2 berechnete Tunnelstromdichte an Elektronen für das „Interface State Injection Model“.

Es wurden diffusionsbegrenzte Kapazitäten verwendet, um mit je einer „Source Measurement Unit“ (SMU) die Ströme am Gate, am Wannenkontakt und am Diffusionskontakt zu messen. Wannenkontakt und Diffusionskontakt wurden jeweils auf Null Volt gegenüber Masse gehalten. Eine genaue Analyse der Stromdichten wird am Beispiel des 3,35 nm dicken

Gateoxides und für den Fall PMOS/GI vorgenommen. Im Gegensatz zu den Abbildungen 4.3, 4.4, 4.5 und 4.9 sind nun die Stromdichten in der Abbildung 4.11 gegenüber der Gatespannung V_g aufgetragen. In der Abbildung 4.11 sind vier experimentelle Beobachtungen zu erkennen, die in der Tabelle 4.2 aufgelistet sind.

I.	$ V_g < 4,75 \text{ V}$	$ J_{\text{Gate}} = J_{\text{Wanne}} $
II.	$ V_g > 3,25 \text{ V}$	$ J_{\text{Diff}} = J_{\text{Tunnel}} $
III.	$ V_g < 3,25 \text{ V}$	Ein anderer Mechanismus führt zu einem Löcherstrom mit umgekehrtem Vorzeichen, was mit einem Vorzeichenwechsel der Stromdichte am Diffusionskontakt J_{Diff} verbunden ist.
IV.	$ V_g > 5 \text{ V}$	$ J_{\text{Wanne}} = J_{\text{Diff}} + J_{\text{Gate}} $ und $ J_{\text{Gate}} = J_{\text{Diff}} = J_{\text{Tunnel}} $

Tab. 4.2: Experimentelle Beobachtungen für die Stromdichten im Fall PMOS/GI. J_{Wanne} , J_{Diff} und J_{Gate} sind die am Wannenkontakt, am Diffusionskontakt und am Gate gemessenen Stromdichten. J_{Tunnel} ist die im Kapitel 4.3.2 berechnete Tunnelstromdichte an Elektronen für das „Interface State Injection Model“.

Im folgenden wird eine Modellvorstellung zur Erklärung dieser Ströme bzw. Stromdichten dargelegt, die anschließend im Vergleich zur Literatur diskutiert wird. Als Ausgangspunkt dient die experimentelle Beobachtung IV in der Tabelle 4.2. Ein Elektron, welches elastisch durch das Oxid tunnelt, verliert im Leitungsband der Anode soviel Energie, daß in der Raumladungszone des n-Substrates ein Elektron in das Leitungsband angehoben werden kann. Dieses addiert sich zum Strom am Wannenkontakt. Das verbleibende Loch liefert einen Beitrag zum Strom am Diffusionskontakt. Damit gilt:

$$|J_{\text{Gate,T}}| \approx |J_{\text{Diff}}| \quad (4.15)$$

$$|J_{\text{Wanne}}| = |J_{\text{Gate,T}}| + |J_{\text{Diff}}| \quad (4.16)$$

$J_{\text{Gate,T}}$ ist die Stromdichte der elastisch, ohne der Unterstützung von Haftstellen getunnelten Elektronen. Die Gleichung 4.15 erklärt die experimentelle Beobachtung II in der Tabelle 4.2. Die Gleichung 4.16 stimmt mit der experimentellen Beobachtung IV in der Tabelle 4.2 überein. Die Abbildung 4.12 enthält eine schematische Darstellung, die einer besseren Veranschaulichung dient. Das elastische Tunneln von Elektronen ist in dieser Darstellung mit der Ziffer 1 markiert und die Generation von Elektron-Loch-Paaren sowie das Anheben von Elektronen in das Leitungsband des Substrates kennzeichnet die Ziffer 2.

Für den „Trap-Assisted Tunneling“-Strom bei kleinen Spannungen können im wesentlichen folgende Ansätze unterschieden werden:

- a) Resonantes Tunneln oder elastisches Tunneln von Elektronen mittels Haftstellen im Oxid [4.17, 4.18, 4.19, 4.20, 4.21].
- b) Inelastisches Tunneln bzw. inelastischer Transport von Elektronen über Haftstellen im Oxid [4.15, 4.22].
- c) Erhöhter Tunnelstrom durch Schottky-Emission [4.23].
- d) Erhöhter Tunnelstrom aufgrund der Verformung der Potentialbarriere durch das Coulombpotential von positiven Ladungen im Oxid.

In den Fällen a), c) und d) hätte das Elektron nach dem Tunnelvorgang genügend potentielle Energie, um ebenfalls den oben beschriebenen Mechanismus zu unterstützen. Dies steht im Widerspruch zu der Messung, die zeigt, daß die Stromdichte am Diffusionskontakt unterhalb von $|V_g| < 4,5 \text{ V}$ nicht der Gatestromdichte folgt und der Betrag der Stromdichte am Wannenkontakt nicht mit der Gleichung 4.16 übereinstimmt. Im Fall b) ist es möglich, daß die bei der Kathode injizierten Elektronen auf ihrem Weg zur Anode Energie verlieren und ihre potentielle Energie nicht ausreicht, um Elektronen vom Valenzband in das Leitungsband des Substrates zu heben. Die Elektronen verlieren ihre Energie entlang eines Pfades aus Haftstellen im Oxid. In der Abbildung 4.12 sind dafür zwei Möglichkeiten skizziert. Beim Mechanismus 3 gelangen die Elektronen in das Leitungsband der Anode und für die Stromdichten gilt:

$$|J_{\text{Gate,TAT}}| \gg |J_{\text{Gate,T}}| \quad (\text{empirische Annahme}) \quad (4.17)$$

$$|J_{\text{Wanne}}| = |J_{\text{Gate,TAT}}| + |J_{\text{Gate,T}}| + |J_{\text{Diff}}| \approx |J_{\text{Gate,TAT}}| \quad (4.18)$$

Dabei ist $J_{\text{Gate,TAT}}$ die durch den inelastischen TAT-Mechanismus verursachte Stromdichte. Die Gleichung 4.18 erklärt die experimentelle Beobachtung I in der Tabelle 4.2. Beim Mechanismus 4 in der Abbildung 4.12 gelangen die Elektronen in das Valenzband der Anode. Dort rekombinieren sie mit Löchern. Über den Diffusionskontakt fließen also Löcher in den Inversionskanal. Ab einer bestimmten Gatespannung überwiegen die Generation von Löchern durch den Mechanismus 2. Der Strom am Diffusionskontakt ändert bei dieser Spannung sein Vorzeichen. Dies erklärt die Beobachtung III in der Tabelle 4.2. In der Abbildung 4.12 sind die inelastischen Transportmechanismen für Elektronen von der Kathode zum Leitungsband und zum Valenzband der Anode separat skizziert, um eine bessere Anschaulichkeit zu erreichen. Realistisch ist der inelastische Transport zwischen Haftstellen mit verschiedenen Energieniveaus, die sich nicht, wie in der Abbildung 4.12 skizziert, in zwei Bereiche trennen lassen.

In der Literatur wird die Ausbeute an Elektron-Loch-Paaren im Substrat als „quantum yield“ bezeichnet. Um „quantum yield“ für Elektronen mit potentiellen Energien kleiner als 5 eV zu messen, wurde von Chang et al. die Gateinjektion von Elektronen bei einem 3,5 nm und einem 4,1 nm dicken Oxid verwendet [4.24]. Die von den Autoren gemessenen Ströme an Gate, Wannenkontakt und Diffusionskontakt sind analog zu den oben beschriebenen Ergebnissen. Als „quantum yield“ wurde von den Autoren das Verhältnis des Stromes am Diffusionskontakt zum Gatestrom gewählt. Bei dieser Definition wird nicht berücksichtigt,

daß es einen inelastischen Anteil im Gatestrom geben kann, der keinen Beitrag zur Generation von Elektron-Loch-Paaren im Substrat liefert. Wie in [4.24] wird diese Meßmethode in der Literatur oft verwendet, um theoretische Berechnungen über die Abnahme von „quantum yield“ mit abnehmender potentieller Energie der Elektronen zu bestätigen. Eine derartige Bestimmung von „quantum yield“ führt bei den experimentellen Ergebnissen in der Abbildung 4.11 für $|V_g| < 4,5 \text{ V}$ zu Werten kleiner als eins. Die sehr gute Übereinstimmung zwischen dem Strom am Diffusionskontakt J_{Diff} und der berechneten Tunnelstromdichte an Elektronen J_{Tunnel} (siehe Abbildung 4.11 und Beobachtung II in Tabelle 4.2) zeigt jedoch, daß der Wert für „quantum yield“ zumindest bis $|V_g| < 3,5 \text{ V}$ nicht kleiner als eins ist.

Aus den gemessenen Stromdichten und der berechneten Tunnelstromdichte an Elektronen ist zu schließen, daß „Trap-Assisted Tunneling“ im wesentlichen auf einen inelastischen Leitungsmechanismus im Oxid zurückzuführen ist und „quantum yield“ in diesem Gatespannungsbereich nicht einfach aus dem Verhältnis von Strom am Diffusionskontakt und Gatestrom bestimmt werden kann.

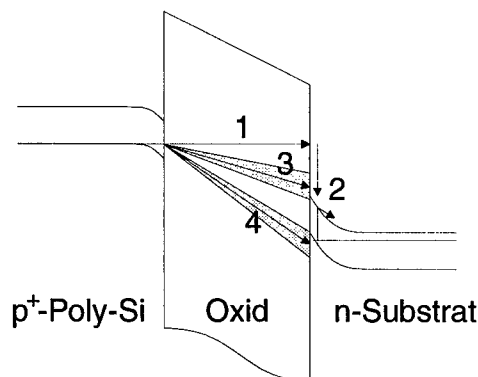


Abb. 4.12: Mechanismen für den Gateleckstrom am Beispiel PMOS/GI. Die mit den Ziffern 1 bis 4 gekennzeichneten Varianten werden im Kapitel 4.4 diskutiert.

4.4.2 „Stress-Induced Leakage Current“

Als „Stress-Induced Leakage Current“ (SILC) wird in der Literatur der zusätzliche Beitrag zum Gateleckstrom im Bereich niedriger Gatespannungen bezeichnet, der nach einem elektrischen Streß des Gateoxides auftritt. Das Verhalten des SILC wird im Kapitel 5 diskutiert, insbesondere die Tatsache, daß ein meßbarer Anstieg im Gateleckstrom erst ab einer Streßspannung von etwa $|V_g| \geq 5 \text{ V}$ zu beobachten ist. Im folgenden soll der Mechanismus für SILC genauer analysiert werden. In der Literatur wird dieser mit einer der

Möglichkeiten für „Trap-Assisted Tunneling“ in Verbindung gebracht, die im Kapitel 4.4.1 genannt wurden.

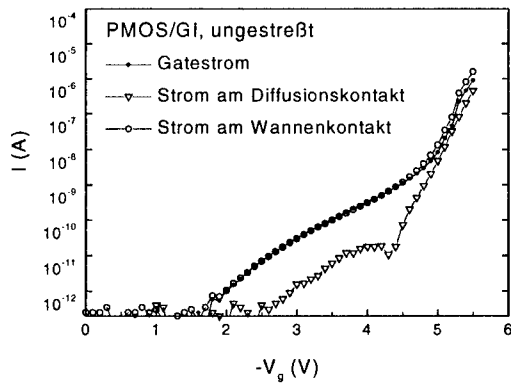


Abb. 4.13: Ströme, gemessen an einer ungestreßten PMOS-Kapazität und für negative Gatespannungen, also für den Fall PMOS/GI. Die Oxiddicke ist 3,35 nm.

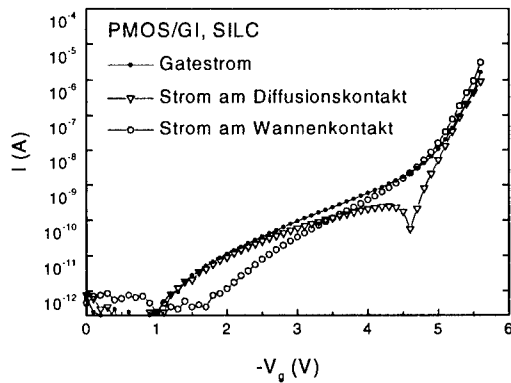


Abb. 4.14: Ströme, gemessen an derselben PMOS-Kapazität wie in der Abbildung 4.13, jedoch nach einem elektrischen Streß bei $V_g = -5,2$ V.

In den Abbildungen 4.13 und 4.14 sind jeweils der Gateleckstrom, der Strom am Diffusionskontakt und der Strom am Wannenkontakt dargestellt. Die Abbildung 4.13 zeigt die Messung an einer ungestreßten PMOS-Kapazität. In der Abbildung 4.14 ist die Messung an derselben Kapazität nach einem elektrischen Streß des Gateoxides dargestellt. Es ist deutlich ein erhöhter Gateleckstrom bei niedrigen Spannungen zu erkennen. Dieser SILC ist durch den Anstieg des Stromes am Diffusionskontakt charakterisiert. Der Strom am Wannenkontakt ändert seinen Verlauf gegenüber der Messung an der ungestreßten Kapazität nicht. Dies läßt vermuten, daß durch den elektrischen Streß der Elektronentransport zwischen Gate und Valenzband erhöht wird. Das entspricht dem Mechanismus 4 in der Abbildung 4.12. Der Elektronentransport zwischen Gate und Leitungsbandkante, welcher in der Abbildung 4.12 mit der Ziffer 3 gekennzeichnet ist, bleibt unverändert. Der Unterschied zwischen den beiden Mechanismen betrifft den Verlust an potentieller Energie der Elektronen beim inelastischen Transport im Oxid. Im ersten Fall ist dieser größer. Die Generation von Haftstellen im Oxid mit einer Ionisationsenergie um die 4 eV könnte dies erklären.

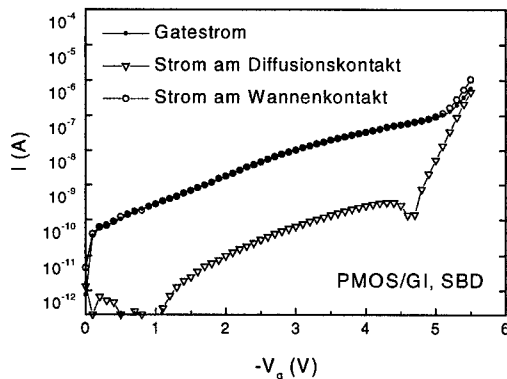


Abb. 4.15: Ströme für dieselbe PMOS-Kapazität wie in der Abbildung 4.13 nachdem sich ein „Soft Breakdown“ ereignet hat.

4.4.3 Leckstrom nach einem „Soft Breakdown“

Für Siliziumdioxidschichten mit einer Dicke kleiner ca. 5 nm wird nach einem elektrischen Streß eine neue Form der Degradation beobachtet, die in der Literatur als „Soft Breakdown“ (SBD) oder „Quasi-breakdown“ bezeichnet wird. Das Kapitel 5 wird sich ausführlicher mit dieser Degradation befassen. Ein Kennzeichen des „Soft Breakdown“ ist, daß der Gateleckstrom bei kleinen Spannungen sprunghaft ansteigt. In der Abbildung 4.15 sind die Ströme nach einem „Soft Breakdown“ gezeigt, die an derselben Kapazität gemessen wurden, wie die Ströme, welche in den Abbildungen 4.13 und 4.14 dargestellt sind. Es ist

festzustellen, daß nach dem SBD der Strom am Wannenkontakt wieder dem Gateleckstrom folgt. Der Strom am Diffusionskontakt bleibt hingegen gegenüber der letzten Messung in der Degradationsphase des SILC unverändert. Der SBD stellt also den stark erhöhten Transport an Elektronen zwischen Gate und Leitungsband dar (Mechanismus 3 in der Abbildung 4.12). Die Generation von Haftstellen im Oxid mit einer Ionisationsenergie um die 3 eV könnte dies erklären.

4.4.4 Auswirkungen auf den PMOSFET

Die in den Kapiteln 4.4.2 und 4.4.3 diskutierten Meßergebnisse zeigen, daß der Gateleckstrom bei kleinen Spannungen, welcher durch „Trap-Assisted Tunneling“ und „Soft Breakdown“ verursacht wird, keinen Einfluß auf den Drainstrom des PMOSFETs hat. Die Ströme erhöhen jeweils nur den Wannenstrom. Im Gegensatz dazu beeinflusst der „Stress-Induced Leakage Current“ den Drainstrom des PMOSFETs.

4.5 Gateleckstrommechanismen für den NMOS und positive Gatespannungen

Für den Fall NMOS/SI wird ein sehr charakteristischer Strom am Wannenkontakt gemessen, der um Größenordnungen kleiner ist als der Gateleckstrom. Das „Valence Band Injection Model“ erklärt diesen Löcherstrom durch das Tunneln von Valenzbandelektronen [4.25]. Als alternative Erklärung wurde ein Modell vorgeschlagen, das diesen Substratstrom durch das Tunneln von Löchern vom Valenzband der Anode in das Valenzband des Substrates erklärt. Diese Löcher werden gemäß dem im Kapitel 2.3.2 erwähnten „Anode Hole Injection Model“ generiert. Diese beiden Modelle werden gegenüber einer Vielzahl weiterer Erklärungsansätze in der Literatur favorisiert. Weinberg und Fischetti geben eine umfassende Zusammenstellung aller Modelle [4.26]. Das Kapitel befaßt sich im folgenden mit dem „Anode Hole Injection Model“ und dem „Valence Band Injection Model“. Im letzteren Fall wird das im Kapitel 4.3 neu eingeführte „Interface State Injection Model“ betrachtet, welches das Tunneln von Valenzbandelektronen für den PMOS und negative Gatespannungen beschreibt.

In der Abbildung 4.16 sind die Ströme für die typischen Phasen der Oxiddegradation zusammengefaßt. Die Messungen wurden jeweils an derselben NMOS-Kapazität mit einer Oxiddicke von 3,8 nm durchgeführt. Zunächst fällt auf, daß der Strom am Diffusionskontakt immer dem Gateleckstrom folgt. Dies ist plausibel, da der Inversionskanal im Fall NMOS/SI die Elektronenquelle für alle Gateleckstrommechanismen darstellt. Damit beeinflussen „Trap-Assisted Tunneling“, „Stress-Induced Leakage Current“ und „Soft Breakdown“ den Drainstrom eines NMOSFETs. Am Wannenkontakt wird während der Degradationsphase des SILC ein entsprechender Anstieg des Stromes gemessen. Nach SBD steigt der Strom am Wannenkontakt bei hohen Spannungen überproportional stark an. Dieser Anstieg kann für das „Anode Hole Injection Model“ nicht mit dem inelastischen Leitungsmechanismus aus dem Kapitel 4.4 bzw. dem Kapitel 4.4.3 erklärt werden. Die Elektronen müssen für die Generation von heißen Löchern ihre potentielle Energie in der Anode verlieren. Ein inelastischer Prozeß im Oxid steht dazu im Widerspruch. Es sei nun angenommen, der Leitungsmechanismus nach

SBD wäre elastisch. Für das „Anode Hole Injection Model“ ist dann die Generation von heißen Löchern in der Anode, welche zur Kathode zurücktunneln würden, proportional zum Elektronenstrom und abhängig von der potentiellen Energie der Elektronen in der Anode. Letztere ist bei gleicher Gatespannung konstant. Damit dürfte der Tunnelstrom an Löchern nach SBD nur um den gleichen Faktor ansteigen, wie der Elektronenstrom. Dies ist jedoch nicht der Fall. Ein anderer Leitungsmechanismus an Löchern im Oxid vom Valenzband des Gates zum Valenzband des Substrates ist nicht erklärbar, da das Polysiliziumgate n^+ -dotiert ist. Die Löcher im Polysiliziumgate müßten dann auf eine andere Weise generiert werden. Es ist daher zu vermuten, daß der am Wannenkontakt gemessene Strom durch Elektronen bzw. „fehlende“ Elektronen, also Löcher im Valenzband des Substrates, entsteht. Bei kleinen Spannungen wird der Elektronentransport durch das Oxid im wesentlichen durch den inelastischen Transport im Oxid verursacht. Dies entspricht in der Abbildung 4.12 den Mechanismen 3 und 4.

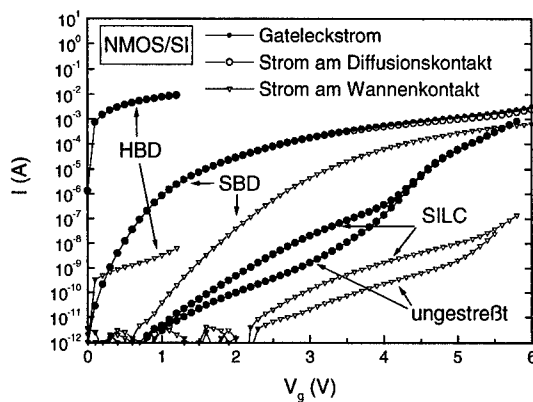


Abb. 4.16: NMOS/Si: Ströme während, bzw. nach den verschiedenen Degradationsphasen. Mit HBD ist der dielektrische Durchbruch des Oxides bezeichnet, der in der Literatur oft „Hard Breakdown“ genannt wird.

Als entscheidendes Argument gegen das „Valence Band Injection Model“ wird in der Literatur angeführt, daß der gemessene Wannenstrom, der meist als Substratstrom bezeichnet wird, zu groß ist. Weinberg und Fischetti finden eine Barrierenhöhe von 3,8 eV bis 3,9 eV [4.26] anstatt von 4,1 eV bis 4,2 eV für die Injektion von Elektronen von der Valenzbandkante. In der Abbildung 4.17 sind neben den Gateleckstromdichten, die bereits in der Abbildung 4.5 gezeigt wurden, auch die am Wannenkontakt gemessenen Stromdichten eingetragen. Die Berechnung der Stromdichten am Wannenkontakt erfolgte mit der Fowler-Nordheim-Gleichung, wobei die Bandverbiegung im Polysilizium und die Gateoxiddicke den Rechnungen für die Gateleckstromdichte entsprechen. Als Barrierenhöhe wurde der Wert

3,9 eV verwendet. Dieses Niveau liegt also zwischen der Valenzbandkante mit 3,1 eV + 1,12 eV \approx 4,2 eV und dem intrinsischen Niveau mit 3,1 eV + 0,56 eV \approx 3,7 eV. Eine Barrierenhöhe von 3,9 eV für die Injektion von Elektronen kann die am Wannenkontakt gemessene Stromdichte erklären.

Das bei der Diskussion des Falles PMOS/GI eingeführte „Interface State Injection Model“ zeigt, daß die Kante des Valenzbandes nicht das obere Energieniveau für Elektronen darstellt, welche effektiv vom Valenzband aus tunneln. Das Energieniveau und damit die Barrierenhöhe ist durch die Grenzflächenzustände gegeben, von welchen aus eine ausreichende Injektion von Elektronen in das Oxid erfolgen kann. Elektronen tunneln vom Valenzband in diese Grenzflächenzustände. Wie bereits im Fall PMOS/GI ist dafür die Barrierenhöhe gleich der Bandlücke E_{gap} . Die Tunneldistanz liegt im Bereich eines Nanometers, wenn das Energieniveau der Valenzbandelektronen nur geringfügig über der Valenzbandkante an der Grenzfläche liegen. Dies erklärt, warum die Barrierenhöhe um ca. 0,2 eV kleiner ist als für die Injektion von der Valenzbandkante. Die Grenzflächenzustände werden bis zu einer Barrierenhöhe von ca. 3,9 eV permanent mit Valenzbandelektronen besetzt. Die im Valenzband verbleibenden Löcher ergeben die am Wannenkontakt gemessene Stromdichte.

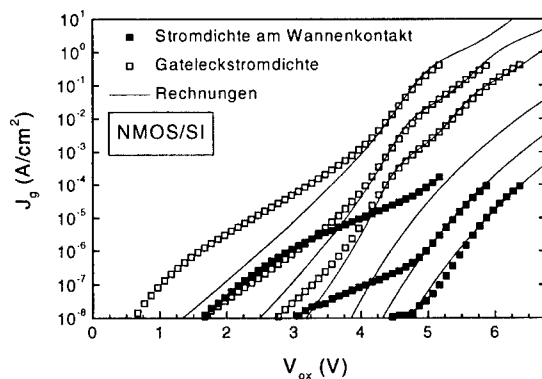


Abb. 4.17: Gateleckstromdichte und Stromdichte am Wannenkontakt für eine ungestreßte NMOS-Kapazität. Die Stromdichten am Wannenkontakt wurden mit der Fowler-Nordheim-Gleichung und einer Barrierenhöhe von 3,9 eV berechnet.

In der Literatur wurde bereits ein Mechanismus diskutiert und als nicht ausreichend beurteilt, welcher eine Art „trap-assisted tunneling“ über Oxiddefekte oder Oberflächenzustände beschreibt [4.26]. Der Unterschied zum hier neu eingeführten „Interface State Injection

Model“ ist, daß von einem zusammenhängenden Tunnelprozeß ausgegangen wurde. Im Kapitel 4.3.1 wurde jedoch aufgezeigt, daß die jeweiligen Grenzflächenzustände permanent besetzt sein können und die Injektion von Elektronen in das Oxid separat betrachtet werden darf.

4.6 Gateleckstrom aufgrund der Verformung der Potentialbarriere durch positive Oxidladungen

In diesem Kapitel wird der Gateleckstrom diskutiert, der nur aufgrund der Verformung der Potentialbarriere durch das Coulombpotential von positiven Ladungen im Oxid entsteht. Es wird untersucht, ob dieser Mechanismus eine Alternative für „Trap-Assisted Tunneling“ ist, um den gemessenen Gateleckstrom eines ungestreßten Oxides erklären zu können.

4.6.1 Potential einer Punktladung im Oxid

Um die Randbedingungen an den Elektroden zu erfüllen, kann die Berechnung des Potentials einer Punktladung im Oxid mit Hilfe der Spiegelladungsmethode erfolgen. In der Abbildung 4.18 ist dies für eine positive Ladung mit den Koordinaten $x = 0$ und $z = z_p$ schematisch skizziert. Die Koordinaten der Punktladung Q_p seien allgemein x_p und z_p . Der Abstand zwischen der Punktladung und einem beliebigen Ort in der x - z -Ebene ist:

$$r_p = \sqrt{(z - z_p)^2 + (x - x_p)^2} \quad (4.19)$$

Spiegelt man die Punktladung zuerst an der Elektrode in negativer z -Richtung, wie in der Abbildung 4.18 gezeigt, dann gilt für die Spiegelladungen:

$$r_j = \sqrt{[z - (-1)^j \cdot (j \cdot T_{ox} + z_p)]^2 + [x - x_p]^2} \quad (4.20)$$

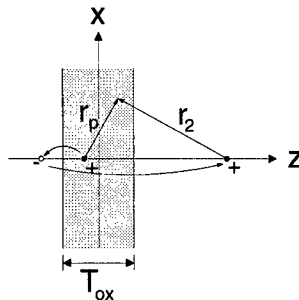


Abb. 4.18: Schematische Darstellung der Potentialberechnung mittels der Spiegelladungsmethode.

Das Potential der Punktladung am Ort $\vec{r}(x, z)$ ist die Summe aller Coulombpotentiale:

$$U_p(x, z) = \frac{Q_p}{4\pi \cdot \epsilon_{\text{ox}}} \left[\frac{1}{r_p} + \sum_{j=1}^{\infty} \frac{(-1)^j}{r_j} \right] \quad (4.21)$$

Bei der Berechnung wird die Reihe abgebrochen, wenn der neue Beitrag zum Potential ein bestimmtes Limit unterschreitet. In der Abbildung 4.19 ist das berechnete Potential für eine positive Punktladung in einem 7 nm dicken Oxid gezeigt. Aus Symmetriegründen ist die Potentialverteilung nur für positive x-Richtungen dargestellt.

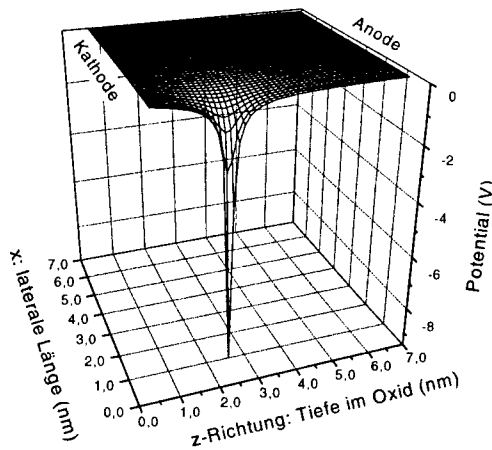


Abb. 4.19: Potential einer positiven Punktladung, die in einer Tiefe von 2,0 nm, d.h. also bei einer z-Koordinate von $z_p = -1,5$ nm lokalisiert ist. In diesem Beispiel liegt keine Potentialdifferenz zwischen der Kathode und der Anode vor. Die Oxiddicke ist 7 nm.

Die Berechnungen, welche in den folgenden Kapiteln diskutiert werden, sind für eine laterale Länge $x = T_{\text{ox}}$ ausgeführt. Für die Berechnung des Tunnelstromes wird jeweils die vereinfachende Annahme gemacht, daß die Elektronen entlang von Geraden senkrecht zur Elektrodenfläche tunneln und nicht entlang von elektrischen Feldlinien. Die Tunnelstromdichte für jede Trajektorie entlang der z-Richtung wird mit dem im Kapitel 4.1.2 beschriebenen Algorithmus berechnet. Die Stromdichte berechnet sich aus dem absoluten Tunnelstrom durch einen Zylinder mit der Kreisfläche $A = T_{\text{ox}}^2 \pi$:

$$J = \frac{1}{x_N^2 \cdot \pi} \cdot \sum_{j=1}^N J_j \cdot (x_j - x_{j-1})^2 \cdot \pi \quad (4.22)$$

Dabei ist $x_N = T_{ox}$, $x_0 = 0$ und J_j die Tunnelstromdichte durch die j -te Trajektorie. Für eine genauere Ortsauflösung in der Nähe der Punktladung wird im Gegensatz zur Abbildung 4.19 eine logarithmische x -Skalierung gewählt.

4.6.2 Positive Ladungen im Oxid

Die Berechnungen zeigen, daß eine positive Ladung, die in der Nähe der Kathode im Oxid lokalisiert ist, den Gateleckstrom im gesamten Spannungsbereich erhöht. Eine Ladung, die sich nahe der Anode befindet, erhöht hingegen den Gateleckstrom nur im Bereich kleiner Spannungen signifikant. Als experimenteller Datensatz zur Anpassung der Modellrechnung wurde die gemessene Tunnelstromdichte für das 3,80 nm dicke Oxid für NMOS/GI gewählt. Die Berechnung der Stromdichte erfolgte zunächst für je eine positive Ladung, die sich bei z -Positionen zwischen 0,5 nm und 1,8 nm befindet (vgl. Abbildung 4.18). In der Abbildung 4.20 ist gezeigt, daß eine Verteilung von positiven Ladungen auf diesen z -Positionen den erhöhten Gateleckstrom bei kleinen Spannungen erklären könnte. Es wurde eine Normalverteilung mit $\sigma = 0,25$ nm und eine Flächenladungsdichte von $1,85 \cdot 10^{10} \text{ cm}^{-2}$ gewählt.

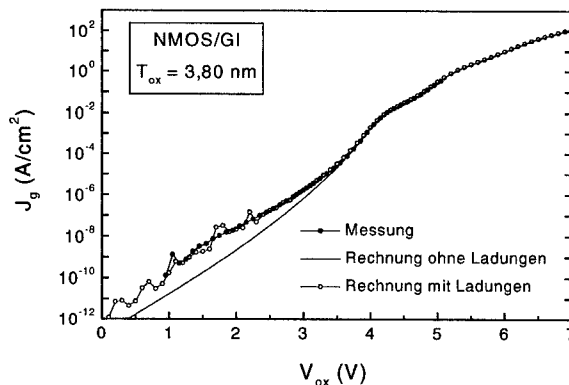


Abb. 4.20: Die erhöhte Stromdichte bei kleinen Spannungen könnte für einen bestimmten Fall mit einer speziellen Verteilung an positiven Ladungen im Oxid erklärt werden. Die gleiche Ladungsverteilung ist aber nicht geeignet für andere Fälle desselben Oxides, wie z.B. NMOS/SI.

Für die im Kapitel 4.1.2 vorgestellte quantenmechanische Berechnung der Tunnelstromdichte treten hier bei bestimmten Trajektorien entlang der z -Richtung Resonanzen auf, die zu einem lokalen Maximum in der Tunnelstromdichte führen. Die Resonanzen sind abhängig von der Weite und der Tiefe des Coulombpotentials an dieser Stelle. Die Resonanzen sind

quantenmechanisch korrekt berechnet, basieren aber auf der Modellannahme, daß die Elektronen parallel zur z-Richtung tunneln. Dadurch erklären sich einzelne Spitzen in der berechneten Stromdichte, die allerdings hier nicht von Bedeutung sind.

Die Größen der verwendeten Parameter sind angepaßt, um für den Fall NMOS/GI eine möglichst gute Beschreibung zu finden. Mit denselben Parametern, also derselben Verteilung an positiven Ladungen, können die Stromdichten für die Fälle NMOS/SI, PMOS/GI und PMOS/SI des gleichen Oxides nicht vergleichbar gut angepaßt werden. Hieraus ist zu schließen, daß der TAT-Strom nicht allein auf die Verformung der Potentialbarriere durch eine Verteilung an positiven Ladungen im Oxid zurückgeführt werden kann. Wäre dies der Fall, dann müßte der TAT-Strom für die ungestreßten Kapazitäten mit jeweils derselben Verteilung an positiven Ladungen beschrieben werden können. Das Tunneln aufgrund der Verformung der Potentialbarriere ist kein inelastischer Prozeß und wäre somit auch im Widerspruch zu der Feststellung im Kapitel 4.4.1, daß der TAT-Strom aufgrund eines inelastischen Transportes von Elektronen im Oxid verursacht wird.

4.7 Auswirkung auf die Zuverlässigkeitsanalyse von Oxiden

Wie im Kapitel 4.2 gezeigt wurde, ist bei Streßexperimenten in Inversion, also NMOS/SI und PMOS/GI, die Bandverbiegung im Polysilizium zu berücksichtigen. Damit hängt die Oxidspannung gemäß Gleichung 4.9 von der Gatespannung ab. In der Abbildung 4.21 sind noch einmal die gemessenen Stromdichten an Elektronen bei hohen Spannungen in einer Fowler-Nordheim-Graphik miteinander verglichen. Der Unterschied für NMOS/SI, aber auch für PMOS/GI, ist deutlich zu erkennen.

Im folgenden wird für die beiden bekannten Modelle, das E-Modell und das I/E-Modell, die Feldbeschleunigung der Oxiddegradation genauer untersucht und u.a. gezeigt, wie sich die Bandverbiegung im Polysilizium auswirkt. Wie bereits im gesamten Kapitel 4 werden die Fälle, die bei Dual-Workfunction-CMOS unterschieden werden müssen, diskutiert. In älteren Technologien wurde für den PMOSFET ein n^+ -dotiertes Polysiliziumgate verwendet. In diesen Technologien tritt also im Fall PMOS/GI keine Bandverbiegung im Polysiliziumgate auf und die Gatespannung und die Oxidspannung unterscheiden sich um einen konstanten Betrag. Das entspricht der folgenden Diskussion für die Akkumulationsfälle.

Zudem wird in der Diskussion zwischen dünnen und dicken Oxiden unterschieden. Der Unterschied ist anhand des Transportes der Elektronen im Oxidleitungsband definiert. Wie bereits im Kapitel 4.2.1 erwähnt wurde, ist der Transport der Elektronen im Oxidleitungsband für Oxiddicken kleiner als ca. 6 nm ballistisch. In dickeren Oxiden verlieren die Elektronen im Oxidleitungsband nach Fowler-Nordheim-Tunneln Energie durch inelastische Stöße. Die Elektronen besitzen jedoch an der Grenzfläche zur Anode eine Verteilung an potentieller Energie gegenüber der Oxidleitungsbandkante, die unabhängig von der Gatespannung oder dem elektrischen Feld im Oxid ist. Dies wird in der Literatur als „electron heating“ bezeichnet [4.27].

Die Unterscheidung zwischen Akkumulation und Inversion bezieht sich immer auf starke Akkumulation und starke Inversion. Die Potentialverhältnisse werden wie im Kapitel 4.2 kalkuliert. Es gelten die dort aufgeführten Gleichungen.

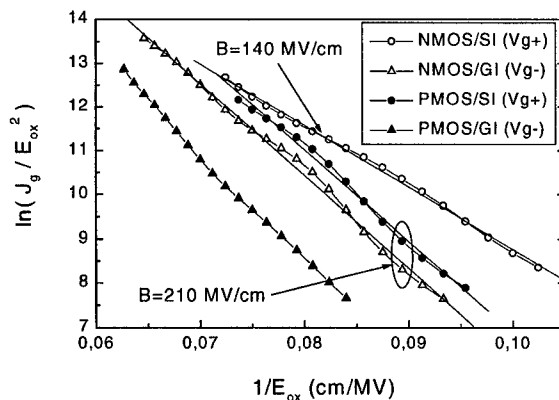


Abb. 4.21: Fowler-Nordheim-Graphik der gemessenen Stromdichten bei hohen elektrischen Feldern. Gezeigt sind die Daten des 4,20 nm dicken Oxides. Die Steigung im Fall PMOS/GI nimmt mit zunehmendem elektrischen Feld im Oxid zu. Für die anderen drei Fälle wurde B durch eine angepaßte Gerade ermittelt.

4.7.1 E-Modell

Akkumulationsfall für dünne Oxide

Im E-Modell ist die primäre Ursache für die Degradation das elektrische Feld im Oxid. Da die Feldbeschleunigung aus der Differenz der elektrischen Felder bestimmt wird, ist für die Akkumulationsfälle eine Umrechnung auf das korrekte elektrische Feld im Oxid nicht notwendig. Beide unterscheiden sich um einen konstanten Betrag. Damit wird in einem Streßexperiment immer die korrekte Beschleunigung der Degradation des Oxides in Abhängigkeit vom elektrischen Feld gemessen. Dies gilt für NMOS/GI ebenso wie für PMOS/SI.

Akkumulationsfall für dicke Oxide

Bei dicken Oxiden kommt es während eines elektrischen Stresses zur Generation von Oxidladungen. Ein Streß bei konstanter Stromdichte regelt die Gatespannung und damit das elektrische Feld nach, so daß das elektrische Feld im Oxid konstant bleibt. Die Feldbeschleunigung kann daher nur aus dem elektrischen Feld im Oxid bestimmt werden.

Inversionsfall für dünne Oxide

Wird für die beiden Inversionsfälle NMOS/SI und PMOS/GI die Bandverbiegung im Polysilizium nicht berücksichtigt, dann ergibt sich eine zu kleine Feldbeschleunigung, weil den ermittelten Durchbruchzeiten zu hohe elektrische Felder zugeordnet werden. Dabei steigt die Differenz mit zunehmendem Gatespannungsbetrag an. Es wird eine Feldbeschleunigung ermittelt, die kleiner ist als für die Akkumulationsfälle und die Extrapolation der Oxiddegradation ergibt eine zu kleine Lebensdauer. Die Feldbeschleunigung ist daher jeweils folgendermaßen zu bestimmen:

$$\gamma = \frac{\ln\left(\frac{t_1}{t_2}\right)}{E_{ox}(E_2) - E_{ox}(E_1)} \quad (4.23)$$

$$E_{ox}(E_g) = \frac{V_{ox}(V_g)}{T_{ox}} \quad (4.24)$$

$$E_g = \frac{V_g}{T_{ox}}$$

$V_{ox}(V_g)$ ist durch die Gleichung 4.9 gegeben. Bei der Extrapolation selbst ist ebenfalls von dem entsprechenden Feld im Oxid auszugehen.

Inversionsfall für dicke Oxide

Prinzipiell gilt hier dieselbe Überlegung wie für dünne Oxide. Sind die Gateoxiddicken jedoch so groß, daß die Streßspannungen in einem Bereich von 10 V oder mehr liegen, dann ist die Korrektur, welche aufgrund der Bandverbiegung im Polysiliziumgate gemacht werden muß, um fast zwei Größenordnungen kleiner als die Streßspannung. Der Fehler, welcher bei der Bestimmung der Feldbeschleunigung unter Vernachlässigung der Korrektur gemacht wird, nimmt also mit zunehmender Oxiddicke ab, da die Streßspannung größer wird.

4.7.2 1/E-Modell

Im 1/E-Modell ist die Ursache der Oxiddegradation der Strom an Löchern, der vom Valenzband der Anode zum Valenzband der Kathode zurücktunnelt. Dieser Löcherstrom ist nach dem „Anode Hole Injection Model“ abhängig von der Stromdichte an Elektronen und der Generation von heißen Löchern in der Anode (vgl. Kapitel 2.3.2). Im 1/E-Modell setzt sich der Beschleunigungsparameter G aus dem Exponentialfaktor B für die Fowler-Nordheim-Tunnelstromdichte an Elektronen und einem Exponentialfaktor H für die Generation der Löcher, die zur Kathode zurücktunneln, zusammen (vgl. Kapitel 2.3.2). Das 1/E-Modell behandelt die Generation von Löchern empirisch. Der Beschleunigungsparameter G wird experimentell ermittelt. Für die Größe H spielt u.a. die potentielle Energie der Elektronen in der Anode eine Rolle. Aufgrund des ballistischen Transportes der Elektronen im Oxidleitungsband wird diese für dünne Oxide spannungsabhängig. Dies wird als Erklärung angegeben, warum der Parameter G im 1/E-Modell für dünne Oxide stark zunimmt [4.28]. Wenn allerdings die potentielle Energie der Elektronen spannungsabhängig wird, dann gilt

dies auch für die Größe H und damit für den Beschleunigungsparameter G . Eine Extrapolation der Lebensdauer auf die Betriebsspannung ist in einem solchen Fall nach dem $1/E$ -Modell mit einem konstanten G nicht mehr exakt. Explizit verringert sich die Größe H mit abnehmender Gatespannung. In einer halblogarithmischen Graphik der Zeit gegenüber $1/E$ wird also die Steigung, welche durch G gegeben ist, mit abnehmender Spannung kleiner. Es wird daher im folgenden untersucht, wie die potentielle Energie der Elektronen von der Spannung abhängt und wie sich dabei die Bandverbiegung im Polysilizium auswirkt. Allgemein gilt, daß im $1/E$ -Modell immer das exakte elektrische Feld im Oxid verwendet werden muß, da dieses die Abhängigkeit der Tunnelstromdichte an Elektronen beschreibt. Dadurch ergibt sich bereits ein Einfluß der Bandverbiegung im Polysilizium.

Für die Definition des Energieverlustes der Elektronen zur Generation heißer Löcher in der Anode gibt es im wesentlichen zwei Möglichkeiten. Die potentielle Energie ist entweder durch die Leitungsbandkante der Anode an der Grenzfläche zum Oxid oder durch das Fermi-niveau der Anode gegeben. Diese beiden Fälle werden im folgenden mit „Leitungsbandkante“ und „Fermi-niveau“ identifiziert. Die Zusammensetzung der potentiellen Energie der Elektronen in der Anode ist in den Abbildungen 4.22 und 4.23 für die Inversionsfälle skizziert. Die Verhältnisse für die Akkumulationsfälle sind einfacher und daher nicht eigens graphisch dargestellt.

Akkumulationsfall für dünne Oxide

Für die Fälle NMOS/GI und PMOS/SI gelten folgende Gleichungen für die potentielle Energie der Elektronen E_{pot} :

$$\text{Leitungsbandkante:} \quad E_{\text{pot}} = q \cdot |V_g| - E_{\text{gap}} = q \cdot |V_{\text{ox}}| + q \cdot \Delta\Phi_b \quad (4.25)$$

$$\text{Fermi-niveau:} \quad E_{\text{pot}} = q \cdot |V_g| \quad (4.26)$$

V_g ist die Gatespannung, E_{gap} ist die Bandlücke von Silizium, V_{ox} ist die Oxidspannung und $\Delta\Phi_b$ ist die Differenz der Barrierenhöhen zwischen Polysiliziumgate/SiO₂ und Siliziumsubstrat/SiO₂ (siehe auch Kapitel 4.2).

Die potentielle Energie ist spannungsabhängig. Damit sind H und G spannungsabhängig. Die Gatespannung und die Oxidspannung unterscheiden sich um einen konstanten Betrag.

Akkumulationsfall für dicke Oxide

Die potentielle Energie der Elektronen gegenüber der Leitungsbandkante im Oxid ist durch „electron heating“ bestimmt. Es ergibt sich eine Verteilung, die unabhängig von der Gatespannung ist [4.27]. Damit gilt für die Fälle NMOS/GI und PMOS/SI:

$$\text{Leitungsbandkante:} \quad E_{\text{pot}} = q \cdot \Phi_b + \Delta E = \text{konst.} \quad (4.27)$$

$$\text{Fermi-niveau:} \quad E_{\text{pot}} = q \cdot \Phi_b + \Delta E + E_{\text{gap}} = \text{konst.} \quad (4.28)$$

Φ_b ist die Barrierenhöhe zwischen Substrat und Oxid bzw. Polysilizium und Oxid. Mit ΔE ist die potentielle Energie eines Elektrons gegenüber der Leitungsbandkante im Oxid bezeichnet.

Da die Verteilung insgesamt unabhängig von der Gatespannung ist, stellt ΔE in diesem Zusammenhang eine konstante Größe dar.

Die potentielle Energie ist nicht spannungsabhängig und die Extrapolation kann mit einem konstanten Beschleunigungsparameter G erfolgen. Die Gatespannung und die Oxidspannung unterscheiden sich um einen konstanten Betrag.

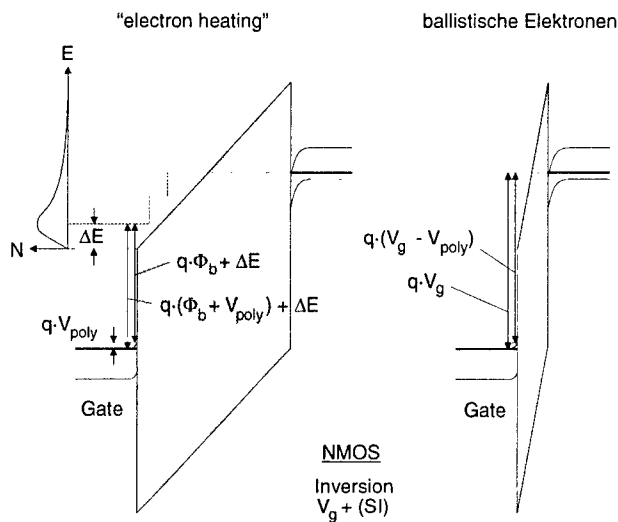


Abb. 4.22: Potentielle Energie der Elektronen in der Anode für den Fall NMOS/Si. Rechts: Es sind die Verhältnisse für dünne Oxide gezeigt, die durch den ballistischen Transport von Elektronen im Oxidleitungsband gekennzeichnet sind. Links: Es sind die Verhältnisse für dicke Oxide gezeigt, die durch „electron heating“ im Oxidleitungsband gekennzeichnet sind.

Inversionsfall für dünne Oxide

In der Abbildung 4.22 sind die Verhältnisse unter der Bezeichnung „ballistische Elektronen“ skizziert. Es gilt für den Fall NMOS/SI:

$$\begin{aligned} \text{Leitungsbandkante: } E_{\text{pot}} &= q \cdot (|V_{\text{p}}| - |V_{\text{poly}}|) = \\ &= q \cdot (|V_{\text{ox}}| - |\Phi_{\text{ms}}| + |V_{\text{Si}}| + \Delta\Phi_{\text{h}}) \end{aligned} \quad (4.29)$$

$$\text{Fermi niveau: } E_{\text{pot}} = q \cdot |V_g| \quad (4.30)$$

V_{poly} ist die Bandverbiegung im Polysiliziumgate, Φ_{ms} ist die Austrittsarbeitsdifferenz zwischen Polysiliziumgate und Substrat und V_{Si} ist die Bandverbiegung im Siliziumsubstrat.

Wenn die Leitungsbandkante maßgebend ist, dann ist die potentielle Energie abhängig von der Oxidspannung. Wenn das Fermi-niveau maßgebend ist, dann ist die potentielle Energie abhängig von der Gatespannung. Der Unterschied zwischen Oxidspannung und Gatespannung ist kein konstanter Betrag sondern wird durch die Bandverbiegung im Polysiliziumgate bestimmt. Der Parameter G ist also spannungsabhängig und es ist das exakte elektrische Feld im Oxid zu verwenden, das nun selbst von der Gatespannung abhängt.

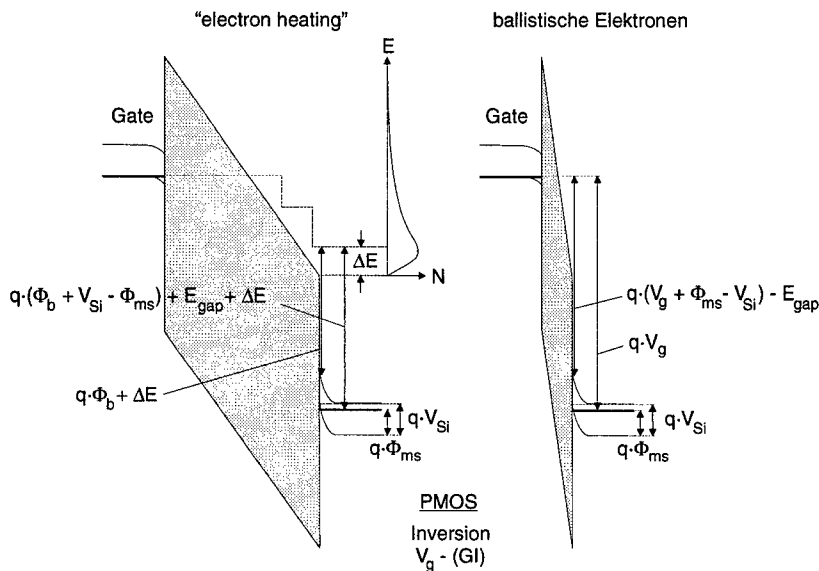


Abb. 4.23: Potentielle Energie der Elektronen in der Anode für den Fall PMOS/GI. Rechts: Es sind die Verhältnisse für dünne Oxide gezeigt, die durch den ballistischen Transport von Elektronen im Oxidleitungsband gekennzeichnet sind. Links: Es sind die Verhältnisse für dicke Oxide gezeigt, die durch „electron heating“ im Oxidleitungsband gekennzeichnet sind.

Für den Fall PMOS/GI gilt (siehe Abbildung 4.23):

$$\text{Leitungsbandkante:} \quad E_{\text{pot}} = q \cdot (|V_g| + |\Phi_{ms}| - |V_{Si}|) - E_{\text{gap}} \quad (4.31)$$

$$\text{Fermi-niveau:} \quad E_{\text{pot}} = q \cdot |V_g| \quad (4.32)$$

Die potentielle Energie ist auch in diesem Fall spannungsabhängig und der Parameter H wird mit abnehmender Gatespannung kleiner. Entsprechend zum Fall NMOS/SI ist das elektrische Feld im Oxid auch abhängig von der Gatespannung. Wie jedoch im Kapitel 4.3.1 gezeigt wurde, ist die Tunnelstromdichte an Elektronen kleiner als im Fall NMOS/SI. Die Tunnelstromdichte kann durch das im Kapitel 4.3.1 eingeführte „Interface State Injection

Model“ beschrieben werden. Der Faktor B ist somit ebenfalls von der Spannung abhängig, da die für die Elektroneninjektion maßgebende Barrierenhöhe spannungsabhängig ist. Dies ist auch in den Abbildungen 4.10 und 4.21 zu erkennen. Der Faktor B wird ebenso wie der Faktor H mit abnehmender Gatespannung kleiner. Insgesamt nimmt also der Beschleunigungsparameter G mit der Gatespannung stärker ab als im Fall NMOS/Si.

Inversionsfall für dicke Oxide

Die Abbildung 4.22 zeigt die Verhältnisse unter der Bezeichnung „electron heating“. Es gilt für den Fall NMOS/Si:

$$\text{Leitungsbandkante:} \quad E_{\text{pot}} = q \cdot \Phi_b + \Delta E = \text{konst.} \quad (4.33)$$

$$\text{Ferminiveau:} \quad E_{\text{pot}} = q \cdot (\Phi_b + V_{\text{poly}}) + \Delta E \quad (4.34)$$

Ist die Leitungsbandkante maßgebend, dann kann ein konstanter Parameter G verwendet werden. Falls das Ferminiveau maßgebend ist, dann wird der Parameter G spannungsabhängig. In diesem Fall variiert die potentielle Energie aber nicht mit der Gate- oder Oxidspannung, sondern nur mit der Bandverbiegung im Polysiliziumgate. Damit ist die Variation eine Größenordnung kleiner als die potentielle Energie selbst. Der Fehler bei einer Vernachlässigung der Abhängigkeit von der Spannung ist wesentlich kleiner als in den oben diskutierten Fällen.

Die elektrische Feldstärke im Oxid ändert sich auch hier mit der Gatespannung. Liegen die Streßspannungen im Bereich von 10 V oder mehr, dann ist die Korrektur, welche aufgrund der Bandverbiegung im Polysiliziumgate gemacht werden muß, um fast zwei Größenordnungen kleiner als die Streßspannung. Wenn diese also vernachlässigt wird, ist der Fehler bei der Ermittlung der Feldbeschleunigung klein.

Im Fall PMOS/GI gilt:

$$\text{Leitungsbandkante:} \quad E_{\text{pot}} = q \cdot \Phi_b + \Delta E = \text{konst.} \quad (4.35)$$

$$\text{Ferminiveau:} \quad E_{\text{pot}} = q \cdot (\Phi_b + |V_{\text{Si}}| - |\Phi_{\text{m}}|) + \Delta E + E_{\text{gap}} \quad (4.36)$$

Damit ist in diesem Fall der Faktor H nicht spannungsabhängig. Für die elektrische Feldstärke gilt entsprechendes wie für den Fall NMOS/Si. Für Dual-Workfunction-CMOS-Technologien ist aber der Faktor B spannungsabhängig, wie bereits für den Fall PMOS/GI und dünne Oxide gezeigt wurde. In Technologien mit einem n^+ -dotierten Polysiliziumgate für den PMOSFET ist dies nicht so, und der Beschleunigungsfaktor G ist in diesem Fall konstant.

4.7.3 Zusammenfassung der Diskussionsergebnisse

In diesem Kapitel wurde untersucht, ob die Verwendung des E-Modells oder des I/E-Modells die wahre Feldbeschleunigung der Oxiddegradation richtig berücksichtigen kann. Für das E-Modell gilt, daß die Feldbeschleunigung in jedem Fall richtig berücksichtigt wird, da hier die Oxiddegradation ausschließlich durch das elektrische Feld im Oxid beschrieben wird. In

Inversion ist allerdings das korrekte elektrische Feld im Oxid zu verwenden, das wegen der Bandverbiegung im Polysiliziumgate abhängig von der Gatespannung ist.

Der Beschleunigungsparameter G des I/E-Modells wird bei dünnen Oxiden generell spannungsabhängig. Dies wird durch eine Spannungsabhängigkeit der potentiellen Energie der Elektronen in der Anode verursacht. Im Inversionsfall PMOS/GI trägt dazu auch die Spannungsabhängigkeit der Barrierenhöhe, die für die Injektion von Elektronen maßgebend ist, bei. Letzteres gilt auch für dicke Oxide. Damit ist für dicke Oxide und im Fall PMOS/GI der Parameter G nur konstant, wenn die CMOS-Technologie ein n^+ -dotiertes Polysiliziumgate verwendet. Eine Spannungsabhängigkeit des Parameters G bedeutet, daß eine Extrapolation mit dem I/E-Modell nicht korrekt ist. Für dicke Oxide und im Fall NMOS/SI variiert die potentielle Energie mit der Bandverbiegung im Polysilizium. Damit ist die Spannungsabhängigkeit des Parameters G geringer. Für die Akkumulationsfälle kann bei dicken Oxiden eine Extrapolation mit einem konstanten Beschleunigungsparameter G erfolgen.

4.8 Zusammenfassung

Im Kapitel 4 wurde gezeigt, daß unter Berücksichtigung der quantenmechanischen Transmission und Reflexion von Elektronen die Gateleckstromdichte durch eine SiO_2 -Schicht bei hohen Spannungen sehr gut berechnet werden kann. Dabei sind für die verschiedenen Fälle, die bei Dual-Workfunction-CMOS unterschieden werden müssen, die korrekten Potentialverhältnisse zu betrachten. Für den PMOS und negative Gatespannungen wurde ein neues Modell eingeführt, das als „Interface State Injection Model“ bezeichnet wurde. Die daraus berechnete Tunnelstromdichte für Elektronen erklärt nicht nur den Gateleckstrom bei hohen Spannungen, sondern auch die Stromdichte am Diffusionskontakt einer diffusionsbegrenzten PMOS-Kapazität oder eines PMOS-Feldeffekttransistors. Aus der Analyse der gemessenen Stromdichten für diesen Fall, konnten Rückschlüsse auf die Mechanismen gewonnen werden, die „Trap-Assisted Tunneling“, „Stress-Induced Leakage Current“ und „Soft Breakdown“ bestimmen. Insbesondere wurde gezeigt, daß „Trap-Assisted Tunneling“ im wesentlichen auf einen inelastischen Transport von Elektronen im Oxid zurückzuführen ist.

Es wurde gezeigt, daß der Strom am Wannenkontakt des NMOS im Fall positiver Gatespannungen durch das Tunneln von Valenzbandelektronen erklärt werden kann. Die Potentialbarriere für das in der Literatur diskutierte „Valence Band Injection Model“ ist hierfür nur geringfügig zu korrigieren. Diese Korrektur wird durch das eingeführte „Interface State Injection Model“ begründet. Für den Prozeß des Tunnelns heißen Löcher, welcher in der Literatur als alternatives Modell sehr oft genannt wird, ergeben sich hingegen Widersprüche.

Die Verformung der Potentialbarriere durch das Coulombpotential von positiven Ladungen im Oxid ist eine denkbare Möglichkeit für die Erklärung des „Trap-Assisted Tunneling“-Stromes bei kleinen Spannungen. Es wurde durch entsprechende Berechnungen für den NMOS und negative Gatespannungen gezeigt, daß eine bestimmte Verteilung an positiven Ladungen im Oxid diese Stromdichte erklären könnte. Dieselbe Ladungsverteilung kann aber

die Stromdichte für den NMOS und positive Spannung sowie für den PMOS nicht beschreiben. Damit kann die Verformung der Potentialbarriere nicht als maßgebende Ursache zur Erklärung des Gateleckstromes aufgrund von „Trap-Assisted Tunneling“ herangezogen werden.

Die im Kapitel 4 erarbeiteten Ergebnisse haben Konsequenzen für die Charakterisierung der Zuverlässigkeit von Gateoxiden. Für dünne Oxide muß im E-Modell das spannungsabhängige elektrische Feld im Oxid berücksichtigt werden, wenn als Streßbedingung der jeweilige Inversionsfall gewählt wird. Es wurde dargelegt, daß der Beschleunigungsparameter G des I/E -Modells für dünne Oxide generell spannungsabhängig wird. Damit ist eine Extrapolation der Lebensdauer zu kleinen Spannungen mit einem konstanten Parameter G nicht exakt. Für dicke Oxide ergibt sich dieses Problem nicht, wenn als Streßbedingung der jeweilige Akkumulationsfall gewählt wird. Für die Inversionsfälle kann der Parameter G als konstant gewählt werden, wenn eine CMOS-Technologie mit einem n^+ -dotierten Polysiliziumgate vorliegt. Bei Dual-Workfunction-Technologien und dicken Oxiden wäre die Spannungsabhängigkeit des Parameters G für den NMOS-Feldeffekttransistor klein. Beim PMOS-Feldeffekttransistor ist sie dann allerdings durch die Spannungsabhängigkeit der Barrierenhöhe nach dem „Interface State Injection Model“ gegeben.

5 Zuverlässigkeitsanalysen an Siliziumdioxidschichten mit Dicken unter 5 nm

Die meßbare elektrische Degradation von Siliziumdioxid ändert sich mit abnehmender Oxiddicke. Für dicke Siliziumdioxidschichten ist ausschließlich der dielektrische Durchbruch zu beobachten. In einem Bereich zwischen ca. 15 nm und 10 nm wird ein sogenannter „Stress-Induced Leakage Current“ (SILC) während des elektrischen Stresses meßbar. Bei Siliziumdioxidschichten dünner als ca. 5 nm schließt sich der SILC-Phase ein neuer Degradationsmodus an, der in der Literatur als „Soft Breakdown“ (SBD) [5.1] oder „Quasi-breakdown“ [5.2, 5.3] bezeichnet wird. Dieser Degradationsmodus ist das zentrale Thema des Kapitels 5.

5.1 Elektrische Degradation

Im folgenden werden die wichtigsten Charakteristika der elektrischen Degradation von Siliziumdioxidschichten zusammengefaßt.

5.1.1 „Stress-Induced Leakage Current“

Im Kapitel 4.4 wurde bereits erwähnt, daß nach elektrischem Streß ein erhöhter Gateleckstrom im Bereich kleiner Spannungen auftritt. Dieser „Stress-Induced Leakage Current“ (SILC) wird durch die Generation von Haftstellen im Oxid verursacht. Der Mechanismus, welcher zu SILC und auch zu einem erhöhten Gateleckstrom am ungestreßten Oxid führt, wird mit „Trap-Assisted Tunneling“ (TAT) bezeichnet und in der Literatur noch kontrovers diskutiert. Im Kapitel 4.4 wurde jedoch gezeigt, daß „Trap-Assisted Tunneling“ im wesentlichen durch inelastisches Tunneln oder einen inelastischen Transport von Elektronen im Oxid verursacht wird. Bei hohen Spannungen bzw. hohen elektrischen Feldern ist dieser Strom kleiner als der Fowler-Nordheim-Tunnelstrom, weshalb sich die Diskussion immer auf den Bereich kleiner Spannungen bezieht.

Der „Stress-Induced Leakage Current“ setzt sich aus zwei Komponenten zusammen [5.4, 5.5]. Eine Gleichstromkomponente, die als „dc SILC“ oder „steady state SILC“ bezeichnet wird, stellt einen zeitunabhängigen Sättigungswert für den Strom dar, der aufgrund von generierten Haftstellen im Oxid fließt. Zudem existiert eine Komponente mit einer 1/-Zeitabhängigkeit, der sogenannte „transient SILC“, welcher durch das Be- oder Entladen von Haftstellen verursacht wird.

Eine weitere wichtige Eigenschaft ist, daß der absolute Betrag des SILC bei einem bestimmten elektrischen Feld mit abnehmender Oxiddicke zunimmt [5.6]. Dies kann durch einen inelastischen Transport von Elektronen im Oxid, nicht aber durch einen elastischen Tunnelprozeß erklärt werden [5.7].

Es existiert ein Limit für den Betrag der Gatespannung, welcher zu SILC führt [5.8, 5.9]. Unterhalb von ca. 5 V wird in einer MOS-Kapazität kein deutlich meßbarer SILC generiert. Diese Eigenschaft kann anhand der Messung mehrerer $I(V_g)$ -Kurven gezeigt werden, wobei die maximale Gatespannung pro Messung sukzessive in Schritten von z.B. 0,2 V erhöht wird. Damit erhält man eine Gruppe von Strom-Spannungs-Kurven. Der bei einer bestimmten Strom-Spannungs-Kurve erkennbare SILC wird durch die Messung aller vorhergehenden Strom-Spannungs-Kurven verursacht. Die Messung der Strom-Spannungs-Kurven stellt an sich den elektrischen Streß dar. In der Abbildung 5.1 ist ein solches Experiment für eine diffusionsbegrenzte NMOS-Kapazität gezeigt. Die Messungen erfolgten für positive Gatespannungen, also für den Fall NMOS/Si. I_{Gate} ist der am Gatekontakt und I_{Wanne} ist der am Wannenkontakt gemessene Strom. Im Kapitel 4.5 wurden bereits die Ströme diskutiert, welche bei diesem Ladungsträgerseparationsexperiment für den NMOS an den verschiedenen Kontakten gemessen werden. Sowohl am Gatekontakt als auch am Wannenkontakt ist der SILC erst für die Strom-Spannungs-Kurve bis zu einer Gatespannung von $V_g = 5,2$ V meßbar. Diese Meßdaten sind gegenüber allen anderen Kurven in der Abbildung 5.1 durch Punkte hervorgehoben. Alle vorher gemessenen Kurven zeigen keinen SILC. Das bedeutet, daß erst die vorhergehende Strom-Spannungs-Kurve bis zur maximalen Gatespannung von $V_g = 5,0$ V eine meßbare Degradation verursacht hat.

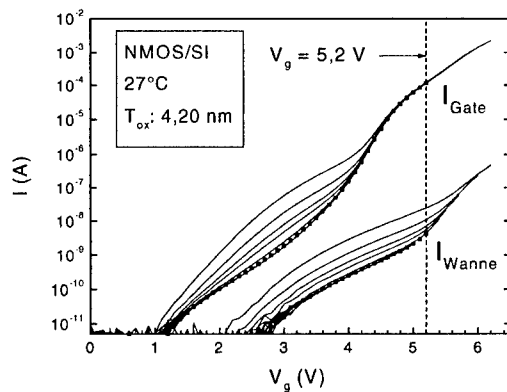


Abb. 5.1: SILC-Degradationsexperiment für den Fall NMOS/Si. Der SILC ist sowohl für den Gateleckstrom I_{Gate} als auch für den Strom am Wannenkontakt I_{Wanne} erst bei der Strom-Spannungs-Kurve bis zu einer Gatespannung von 5,2 V meßbar. Diese Daten sind durch Punkte markiert.

In der Abbildung 5.2 ist das entsprechende Experiment für eine diffusionsbegrenzte PMOS-Kapazität bei negativen Gatespannungen gezeigt, also für den Fall PMOS/GI. Der SILC ist hier ebenfalls erst für die Strom-Spannungs-Kurve bis zu einer Gatespannung von $|V_g| = 5,2$ V meßbar. Diese Meßdaten sind wieder durch Punkte gegenüber allen anderen Strom-Spannungs-Kurven in der Abbildung 5.2 hervorgehoben. Der Anstieg des SILC ist besonders gut für den Strom am Diffusionskontakt I_{Diff} zu erkennen. Im Kapitel 4.4 wurde bereits ein Modell zur Erklärung der Gateleckstrommechanismen im Fall PMOS/GI entwickelt, das den unmittelbaren Zusammenhang zwischen SILC und der Stromdichte am Diffusionskontakt erklärt.

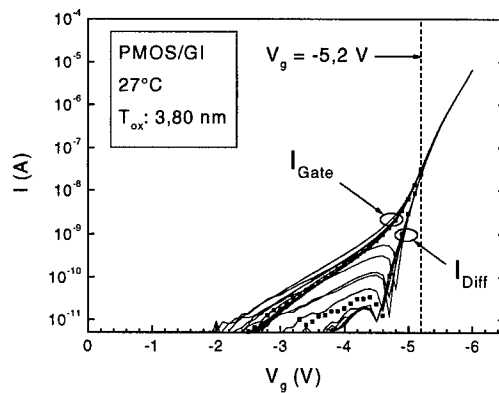


Abb. 5.2: SILC-Degradationsexperiment für den Fall PMOS/GI. Der SILC ist sowohl für den Gateleckstrom I_{Gate} als auch für den Strom am Diffusionskontakt I_{Diff} erst bei der Strom-Spannungs-Kurve bis zu einer Gatespannung von 5,2 V meßbar. Diese Daten sind durch Punkte markiert.

Entsprechende Experimente, die für den Fall NMOS/SI mit unterschiedlichen Oxiddicken im Bereich von 3,8 nm bis 4,7 nm durchgeführt wurden, bestätigen jeweils die 5 V-Grenze. Somit ist nicht das elektrische Feld im Oxid für die Degradation verantwortlich, die SILC verursacht. Die Degradation wird durch die potentielle Energie der Elektronen bestimmt [5.8, 5.9]. DiMaria et al. bringen die Generation von SILC in Zusammenhang mit einem von den gleichen Autoren vorgestellten Modell für die Generation von Haftstellen im Oxid („Trap Creation Model“) [5.10]. Das darin enthaltene „Hydrogen Release Model“ kann die 5 V-Grenze erklären [5.9]. Die potentielle Energie der Elektronen ist bei Oxiden dünner als ca. 6 nm abhängig von der Gatespannung, was bereits bei der Diskussion im Kapitel 4.7 behandelt wurde. Ab einer Gatespannung von 5 V besitzen die Elektronen eine potentielle Energie bezüglich der Oxidleitungsbandkante von mindestens 2 eV und können damit nahe der Anode Wasserstoff im Oxid freisetzen. Dieser generiert neutrale Haftstellen für

Elektronen im Oxid („neutral electron traps“). DiMaria et al. zeigten den Zusammenhang zwischen der Generation von SILC mit der Generation von neutralen Haftstellen für Elektronen im Oxid [5.9].

5.1.2 „Soft Breakdown“

Charakteristika des „Soft Breakdown“

Die als „Soft Breakdown“ (SBD) oder „Quasi-breakdown“ bezeichnete Degradation tritt im Gegensatz zum „Stress-Induced Leakage Current“ (SILC) lokal auf [5.2, 5.11]. Damit ist die Bezeichnung „Durchbruch“ in Anlehnung an den dielektrischen Durchbruch des Oxides gerechtfertigt. Der SBD schließt sich der Degradationsphase des SILC an und führt zu einem Gateleckstrom im Bereich kleiner Spannungen, der deutlich größer ist als der zuletzt gemessene SILC. Somit ist der SBD durch einen diskontinuierlichen Anstieg im Gateleckstrom bei kleinen Spannungen gekennzeichnet. Dies kann genutzt werden, um während eines Streßexperimentes den Zeitpunkt des SBD zu bestimmen [5.12]. Die Meßroutine wird im nächsten Abschnitt beschrieben. Der diskontinuierliche Anstieg im Gateleckstrom entspricht bei einem Streß mit konstanter Stromdichte einem plötzlichen Spannungsabfall. Die abrupte Änderung des Gatesignals ist um so ausgeprägter, je kleiner die Fläche der MOS-Kapazität ist [5.13]. Dies ist auch ein Hinweis auf den lokalen Charakter des SBD. Der Gateleckstrom, welcher nach einem SBD gemessen wird, hat im Bereich kleiner Spannungen eine ähnliche $I_g(V_g)$ -Charakteristik wie der SILC. In der Literatur findet man daher die Bezeichnung „B-mode SILC“ [5.14].

Das Gatesignal weist nach SBD ein starkes Rauschen auf. In der Literatur werden auch Fluktuationen zwischen zwei oder mehreren diskreten Niveaus berichtet [5.13, 5.15, 5.16] und daher oft die Bezeichnung „random telegraph signal“ gewählt. Als mögliche Erklärung für diese Fluktuationen wird das Öffnen oder Schließen eines Durchbruchspfades zwischen Anode und Kathode durch eine Haftstelle im Oxid herangezogen [5.15, 5.17]. Der Zeitpunkt des SBD kann daher während eines Streßexperimentes auch durch den Anstieg des Rauschens im Gatesignal bestimmt werden [5.18].

Die Degradation des Oxides nach dem ersten SBD ist durch das Auftreten weiterer SBD-Leckstrompfade gekennzeichnet. Der Ort des dielektrischen Durchbruches ist unabhängig vom Ort der vorher aufgetretenen SBD-Pfade [5.11].

Meßroutine zur Bestimmung des „Soft Breakdown“

Während eines Streßexperimentes kann die Zeit bis zum SBD durch den diskontinuierlichen Anstieg des Gateleckstromes bestimmt werden [5.12]. Im folgenden wird von einem elektrischen Streß bei konstanter Spannung ausgegangen. Ist der Fowler-Norheim-Tunnelstrom bei der gewählten Streßspannung größer als der Gateleckstrom nach SBD, dann muß zusätzlich der Gateleckstrom bei einer kleineren Spannung betrachtet werden. In logarithmischen Zeitabständen wird daher der elektrische Streß kurz unterbrochen, um den Strom bei einer kleineren Spannung zu messen. Diese wird im folgenden als Monitorspannung und der entsprechende Gateleckstrom als Monitorstrom bezeichnet. In der

Abbildung 5.3 sind die $J_g(V_g)$ -Kurven während bzw. nach den jeweiligen Degradationsphasen gezeigt. Zudem sind exemplarisch die Streßspannung und die Monitorspannung eingezeichnet. In der Abbildung 5.4 ist ein Beispiel für einen Datensatz dargestellt, der während eines elektrischen Stresses mit dieser Meßroutine erzielt wird. Nach dem ersten SBD folgt die Degradationsphase mit weiteren SBD-Ereignissen bis zum dielektrischen Durchbruch.

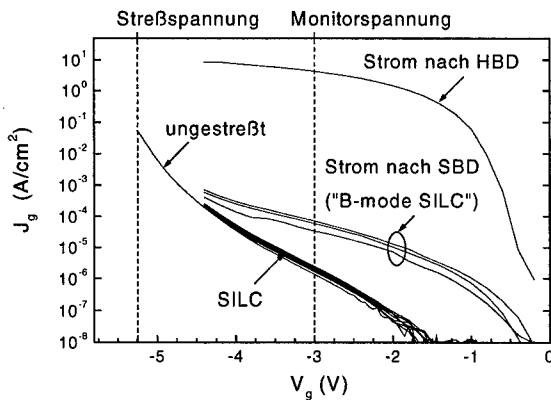


Abb. 5.3: Beispiel für die Gateleckströme während bzw. nach den verschiedenen Degradationsphasen „Stress-Induced Leakage Current“ (SILC), „Soft Breakdown“ (SBD) und „Hard Breakdown“ (HBD). Die Streßspannung und die Monitorspannung für die Meßroutine zur Bestimmung der Zeit bis zum SBD sind exemplarisch eingezeichnet.

Die kurze Unterbrechung des elektrischen Stresses bei konstanter Spannung beeinflusst nicht die gemessene Zeit bis zum dielektrischen Durchbruch, dem „Hard Breakdown“ (HBD) [5.19]. Die Zeit bis zum SBD wird mit einer konstanten relativen Genauigkeit bestimmt, wenn die Messung des Monitorstromes in logarithmisch ansteigenden Zeitabständen erfolgt:

$$t_n = t_{n-1} \cdot 10^a \quad a = \text{konst.} \quad (5.1)$$

Dabei ist t_n der Zeitpunkt der n-ten Messung des Monitorstromes. Für die relative Genauigkeit der gemessenen Zeit bis zum SBD gilt dann:

$$\frac{\Delta t_n}{t_n} = \frac{t_n - t_{n-1}}{t_n} = 1 - 10^{-a} \quad (5.2)$$

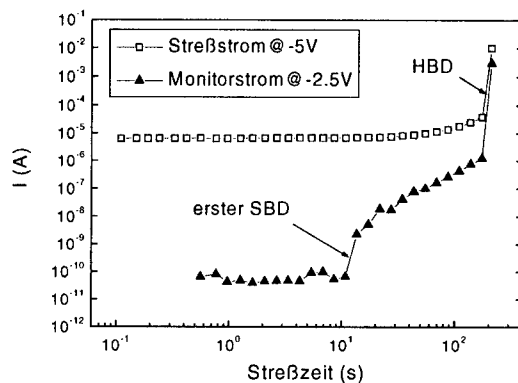


Abb. 5.4: Monitorstrom und Streßstrom während eines elektrischen Stresses mit konstanter Spannung an einer MOS-Kapazität. Der erste „Soft Breakdown“ (SBD) wird mit Hilfe des Monitorstromes gemessen. In der weiteren Degradationsphase bis zum dielektrischen Durchbruch („Hard Breakdown“, HBD) finden weitere SBD-Ereignisse statt.

5.1.3 Der dielektrische Durchbruch – „Hard Breakdown“

Die genauen Mechanismen, welche zum dielektrischen Durchbruch von Siliziumdioxid führen, sind noch nicht endgültig geklärt. In der Literatur gilt jedoch generell als akzeptiert, daß der dielektrische Durchbruch in zwei Phasen verläuft [5.20, 5.21]. Zunächst findet eine längere Phase der Degradation statt, in der Oxiddefekte oder Haftstellen generiert werden. Der eigentliche dielektrische Durchbruch wird durch einen schnellen, lokalen thermischen Prozeß verursacht, der zu einem leitenden Filament im Oxid führt. Dies findet dort im Oxid statt, wo die Defekte oder Haftstellen entweder einen sehr hohen lokalen Strom ermöglichen oder die Bindungsstruktur zerstören.

Der dielektrische Durchbruch, oder auch „Hard Breakdown“ (HBD) genannt, findet lokal statt. Existiert innerhalb der gestreßten Gatefläche einer MOS-Kapazität ein Bereich im Oxid, der bereits vor dem Streß Defekte aufweist, dann erfolgt der HBD zuerst an dieser schwächsten Stelle im Oxid. Streßexperimente werden üblicherweise nach dem Auftreten eines HBD abgebrochen. Mit Hilfe von Emissionsspektroskopieanalysen wird im folgenden gezeigt, daß nach dem ersten HBD weitere HBD-Pfade im Oxid gebildet werden, wenn das elektrische Feld erhöht wird. In der Abbildung 5.5 sind zwei Analysen derselben MOS-Kapazität gezeigt. Der photographischen Aufnahme der Waferoberfläche ist jeweils die Analyse der Photoemission überlagert. Die erste Analyse erfolgte, nachdem ein HBD detektiert wurde. Der Ort des HBD-Pfades ist durch einen Punkt gekennzeichnet, dessen

Durchmesser ein Maß für die Photoemission von Ladungsträgern in diesem HBD-Pfad ist. Die zweite Analyse wurde vorgenommen, nachdem die Streßspannung und damit das elektrische Feld gegenüber dem ursprünglichen Wert erhöht wurde. Es sind mehrere neue HBD-Pfade, verteilt über die gesamte Gatefläche der MOS-Kapazität, zu erkennen.

Für Oxiddicken kleiner als ca. 5 nm ist festzustellen, daß in der Weibull-Verteilung die Steigung der intrinsischen Fehlergeraden mit abnehmender Oxiddicke kleiner wird. Dies kann erklärt werden, wenn der Durchbruch des Oxides auf eine kritische Dichte an Defekten oder Haftstellen zurückgeführt wird. Im „Percolation Model“ für den intrinsischen Durchbruch ist diese kritische Dichte erreicht, wenn ein leitender Pfad aus Haftstellen zwischen den Elektroden vorliegt [5.22]. Den Haftstellen wird eine Sphäre zugeordnet, deren Überlapp entscheidend für eine leitfähige Verbindung zwischen zwei Haftstellen ist. Degraeve et al. konnten die Steigung der intrinsischen Fehlergeraden, den sogenannten „Weibull slope“, mit Monte-Carlo-Simulationen sehr gut berechnen, wenn für den Sphärenradius der Wert 0,45 nm verwendet wurde [5.22, 5.23]. Damit wurde auch endgültig gezeigt, daß der dielektrische Durchbruch von SiO_2 keiner intrinsischen Materialeigenschaft entspricht, sondern ein statistisches Phänomen ist. Die Bezeichnung „intrinsischer Durchbruch“ ist historisch bedingt. In den gleichen Arbeiten werden Argumente dargelegt, daß die Degradation, welche zum dielektrischen Durchbruch führt, im Zusammenhang mit der Generation von neutralen Haftstellen für Elektronen im Oxid steht.

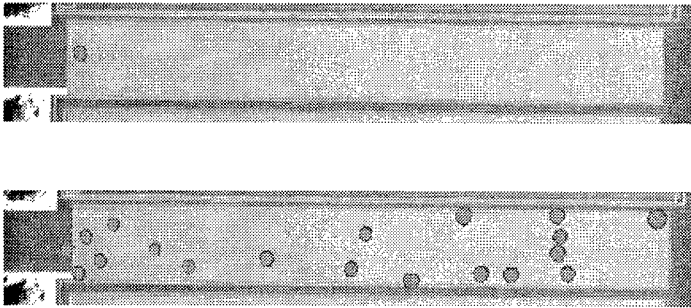


Abb. 5.5: Oben: Emissionsspektroskopie an einer MOS-Kapazität nach „Hard Breakdown“ (HBD). Der HBD ist am linken Rand lokalisiert. Unten: Emissionsspektroskopie an derselben MOS-Kapazität, nachdem das elektrische Feld weiter erhöht wurde. Es sind mehrere zusätzliche HBD-Pfade im Oxid aufgetreten.

Nach dem „Percolation Model“ wird die kritische Haftstellendichte mit abnehmender Oxiddicke kleiner, da weniger Haftstellen pro Gatefläche ausreichen, um einen leitenden Durchbruchspfad zwischen den Elektroden zu bilden. Dies führt in der Literatur zu Aussagen über die Zuverlässigkeit sehr dünner Oxide und zur Vorhersage eines Skalierungslimits [5.24].

Die Tatsache, daß der „Weibull slope“ β mit abnehmender Oxiddicke kleiner wird, hat zur Folge, daß die Abhängigkeit der Zeit bis zum Durchbruch t_{BD} von der Gatefläche zunimmt. Dies resultiert aus dem im Kapitel 2.3.4 diskutierten Zusammenhang zwischen t_{BD} und der Fläche A .

5.1.4 Elektrischer Streß mit konstanter Spannung

Bei dünnen Oxiden wird ein Streß bei konstanter Spannung gewählt, im Gegensatz zu einem Streß bei konstanter Stromdichte im Fall von dicken Oxiden. In der Literatur wird als Grund angegeben, daß sich mit abnehmender Oxiddicke der Transport von Elektronen im Oxidleitungsband ändert [5.25]. Diese Änderung wurde bereits in der Diskussion im Kapitel 4.7 behandelt. Sie ist nicht die korrekte Ursache für den Wechsel der Streßbedingung, spielt aber bei der Analyse der Oxiddegradation eine Rolle. Daher wird im folgenden die Thematik exakt formuliert.

Beim Fowler-Nordheim-Tunneln gelangen Elektronen in das Leitungsband des Oxides und gewinnen dort im elektrischen Feld Energie. In Oxiden dicker als ca. 6 nm verlieren sie diese wieder durch inelastische Stoßprozesse, z.B. mit Phononen. Der Prozeß aus Energiegewinn und Energieverlust ergibt eine Verteilung der potentiellen Energie der Elektronen gegenüber der Leitungsbandkante des Oxides [5.26]. Unabhängig vom elektrischen Feld im Oxid liegt das Maximum der Stromdichte an Elektronen im Bereich von wenigen Elektronenvolt. Dieser Vorgang wird in der Literatur als „electron heating“ im Oxid bezeichnet und ist erst seit Mitte der 80er Jahre durch obiges Modell erklärbar. Die Elektronen besitzen also eine potentielle Energie gegenüber dem Fermi-niveau der Anode, die sich aus der Barrierenhöhe zwischen Silizium und SiO_2 , sowie der potentiellen Energie gegenüber der Leitungsbandkante des Oxides zusammensetzt. Letztere ist für die Elektronen ebenso unabhängig von der Gatespannung wie die Barrierenhöhe. Nur die maximal mögliche potentielle Energie gegenüber der Leitungsbandkante des Oxides ist vom elektrischen Feld im Oxid und damit von der Gatespannung abhängig.

Mit abnehmender Oxiddicke wird der Transport der Elektronen im Oxidleitungsband ballistisch [5.27]. Wie bereits im Kapitel 4.2.1 diskutiert, äußert sich das für Oxide dünner als ca. 6 nm in Oszillationen im Fowler-Nordheim-Tunnelstrom. Die Elektronen verlieren damit keine Energie durch inelastische Stöße im Oxid und ihre potentielle Energie gegenüber dem Fermi-niveau in der Anode ist durch die Gatespannung bestimmt.

Ausgehend von diesem Unterschied ist nun die Oxiddegradation in Abhängigkeit von den bekannten Degradationsmodellen zu betrachten. Es existieren in der Literatur zwei Modelle, welche die Oxiddegradation in einen Zusammenhang mit der potentiellen Energie der Elektronen bringen. Dies sind das „Anode Hole Injection Model“, welches im Kapitel 2.3.2 bereits im Zusammenhang mit dem $1/E$ -Modell erklärt wurde, und das „Hydrogen Release Model“ [5.10]. Bei dicken Oxiden führt die Generation von Oxidladungen zu einer Verkleinerung des elektrischen Feldes im Oxid. Aus diesem Grund wird ein Streß mit konstanter Stromdichte gewählt, der die Streßspannung nachregelt, um das elektrische Feld im Oxid konstant zu halten. Für einen Streß mit konstanter Stromdichte ergibt sich eine gleichbleibende Degradation dicker Oxide über die gesamte Streßzeit, da gilt:

- a) Die Flußdichte an Elektronen, die in diesen Modellen die Ursache der Degradation ist, bleibt mit der Stromdichte konstant.
- b) Eine Veränderung der Gatespannung ändert nicht die potentielle Energie der Elektronen in der Anode oder im Oxid.

Für dünne Oxide tritt keine wesentliche Generation von Oxidladungen auf. Das elektrische Feld im Oxid wird nicht kleiner und bei einem Streß mit konstanter Spannung bleibt die Stromdichte konstant. Zudem wird die potentielle Energie der Elektronen in der Anode oder im Oxid durch die Gatespannung bestimmt. Damit ist für dünne Oxide während eines elektrischen Stresses bei konstanter Spannung eine gleichbleibende Degradation gegeben. Da das elektrische Feld im Oxid bei einem solchen Streß auch konstant bleibt, ändert sich die Fowler-Nordheim-Tunnelstromdichte nicht. Je nach Oxiddicke gibt es aber einen Spannungsbereich, in dem die Stromdichte nach einem „Soft Breakdown“ größer als die Fowler-Nordheim-Tunnelstromdichte ist. Mit abnehmender Oxiddicke liegt die Streßspannung immer mehr im Bereich des direkten Tunnelns. Damit ändert sich die Stromdichte aufgrund des „Stress-Induced Leakage Current“. Im Kapitel 4.4 wurde gezeigt, daß die Elektronen sowohl beim „Stress-Induced Leakage Current“ als auch beim Strom nach einem „Soft Breakdown“ nicht die potentielle Energie besitzen, um zur Degradation nach den obengenannten Modellen beitragen zu können. Damit würde ein Streß bei konstanter Stromdichte nur zu einer gleichbleibenden Degradation führen, wenn über die gesamte Streßzeit ausschließlich die Fowler-Nordheim-Tunnelstromdichte gemessen wird. Für dünne Oxide wird daher ein Streß bei konstanter Spannung bevorzugt.

Nach dem E-Modell ist die Degradation des Oxides nicht durch die potentielle Energie der Elektronen bestimmt, sondern durch das elektrische Feld im Oxid. Wie die Diskussion gerade gezeigt hat, bleibt dieses bei dicken Oxiden für einen Streß mit konstanter Stromdichte unverändert. Bei dünnen Oxiden ist das elektrische Feld während eines elektrischen Stresses mit konstanter Spannung ebenfalls konstant. Somit gilt die Aussage bezüglich der Streßbedingungen auch, falls eine Form der Oxiddegradation durch das E-Modell korrekt beschrieben wird.

Der Wechsel von einem Streß mit konstanter Stromdichte zu einem Streß mit konstanter Spannung ist ab Oxiddicken kleiner ca. 5 nm sinnvoll. Der Grund hierfür ist, daß mit abnehmender Oxiddicke die gemessene Stromdichte bei einem elektrischen Streß immer mehr durch „Stress-Induced Leakage Current“ und „Soft Breakdown“ beeinflusst wird. Ein Streß bei konstanter Spannung ist zulässig, weil:

- der Transport von Elektronen im Oxidleitungsband überwiegend ballistisch ist.
- keine wesentliche Generation von Oxidladungen und damit keine Änderung des elektrischen Feldes im Oxid stattfindet. Die Fowler-Nordheim-Tunnelstromdichte und die direkte Tunnelstromdichte nehmen also bei einem Streß mit konstanter Spannung nicht mehr ab.

5.2 Einfluß von „Soft Breakdown“ auf MOS-Feldeffekttransistoren

5.2.1 Gateleckstrom und Signalrauschen

Aufgrund der im Kapitel 5.1.2 beschriebenen Charakteristika des „Soft Breakdown“ (SBD) müssen der Gateleckstrom und das Rauschen im Gatesignal diskutiert werden. Der erhöhte Gateleckstrom im Bereich kleiner Gatespannungen ist um Größenordnungen kleiner als der übliche Drainstrom, so daß keine meßbare Änderung des Drainstromes auftritt. Der Gateleckstrom führt aber zu einer höheren Verlustleistung pro aktive Gatefläche.

Bei dicken Oxiden wird das 1/f-Rauschen im Drainstrom eines MOSFETs durch das Einfangen und Freisetzen von Elektronen in Haftstellen verursacht. Damit ist dieses Rauschen nicht mit einem 1/f-Rauschen am Gate korreliert. Der Austausch von Ladungsträgern zwischen Kanal und Gate ist bei dünnen Oxiden größer und beeinflusst den Drainstrom. Der Einfluß von „Stress-Induced Leakage Current“ und „Soft Breakdown“ auf den Inversionskanal von NMOSFET und PMOSFET wurde bereits in den Kapiteln 4.4.4 und 4.5 diskutiert. Der Beitrag zum 1/f-Rauschen des Drainstromes, der insbesondere nach einem „Soft Breakdown“ auftritt, ist daher korreliert mit dem 1/f-Rauschen am Gate. Weir et al. fanden jedoch, daß dieser Beitrag nach „Soft Breakdown“ vernachlässigbar klein gegenüber dem 1/f-Rauschen im Drainstrom ist, das nicht mit dem Gatesignal korreliert ist [5.18]. Es ist dennoch sinnvoll, den Einfluß des „Soft Breakdown“ mit einzubeziehen, wenn das 1/f-Rauschen für eine bestimmte Schaltung wichtig ist.

Außer diesen Charakteristika, die bisher nur an MOS-Kapazitätsstrukturen analysiert wurden, waren keine Einflüsse auf MOS-Feldeffekttransistoren bekannt. Die folgenden Kapitel befassen sich mit den ersten Streßexperimenten an MOS-Feldeffekttransistoren mit dünnen Gateoxiden und den daraus gewonnen Ergebnissen [5.12].

5.2.2 Streßexperimente an MOS-Feldeffekttransistoren

Die Streßexperimente an MOS-Feldeffekttransistoren zur Analyse des SBD verwenden im wesentlichen dieselbe Streßroutine, die im Kapitel 5.1.2 vorgestellt wurde. Die Diffusionskontakte und der Wannenkontakt des Transistors werden während des elektrischen Stresses des Gateoxides auf Null Volt gegenüber Masse gehalten. Der Streß erfolgt bei 125°C, um die Oxiddegradation zu beschleunigen. Die Unterbrechung des Stresses für die Messung des Monitorstromes wird gleichzeitig genutzt, um die wichtigen Parameter des Transistors zu bestimmen. Dies ist möglich, wenn Source und Drain jeweils mit einer „Source Measurement Unit“ (SMU) kontaktiert sind. Die ermittelten Transistorparameter sind zum einen die Einsatzspannungen und die Sättigungsdrainströme für $|V_{\text{G}}| = 0,1 \text{ V}$ und für $|V_{\text{G}}| = V_{\text{DD}}$. V_{DD} ist die jeweilige Versorgungsspannung, die technologiespezifisch ist. Zum anderen werden der Off-Strom, die Steigung der Unterschwellgeraden, sowie das Maximum der Transferleitfähigkeit bestimmt. Ein Überblick über die verschiedenen Parameter ist im Kapitel 7.2 des Anhangs A enthalten. Die Streßexperimente werden je nach Wahl abgebrochen, wenn ein SBD oder ein HBD im Gateoxid aufgetreten ist. Zusätzlich werden

die Eingangskennlinien des MOSFETs vor Beginn und nach Abbruch des Streßexperimentes aufgezeichnet.

5.2.3 „Gate-Induced Drain Leakage“ bei NMOSFETs

Im Fall von NMOS-Kurzkanaltransistoren ist die einzige meßbare Degradation, die nach einem SBD im Gateoxid auftritt, durch eine Erhöhung des Off-Stromes gegeben [5.12]. Dabei fanden sich in einem konkreten Streßexperiment nur ein kleiner Anteil von ca. 10% aller Transistoren mit einer Gatelänge von 0,20 µm, welche diesen Effekt aufwiesen. Alle gestreßten NMOSFETs mit einer Gatelänge von 10 µm zeigten diesen Effekt nicht. Wie im weiteren dargelegt wird, ist die Erhöhung des Off-Stromes auf eine Erhöhung des „Gate-Induced Drain Leakage“ (GIDL) zurückzuführen [5.12]. Im Anhang A befaßt sich das Kapitel 7.5 mit dem wesentlichen Mechanismus, der dem GIDL zugrunde liegt.

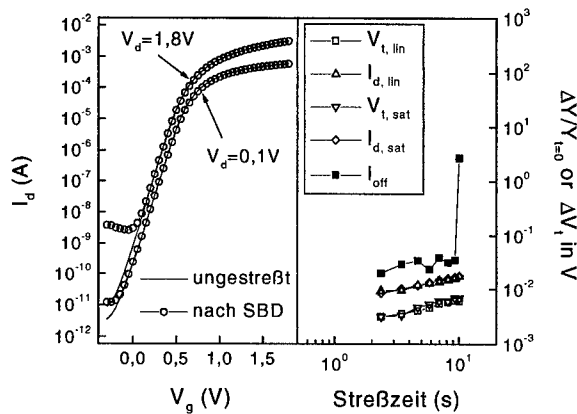


Abb. 5.6: NMOS-Kurzkanaltransistor, $W/L = 10/0,20 \mu\text{m}$, der einen erhöhten Off-Strom bei $V_d = V_{DD} = 1,8 \text{ V}$ nach „Soft Breakdown“ (SBD) aufweist. Im linken Teil sind die Eingangskennlinien vor Beginn des Stresses und nach SBD gezeigt. Der rechte Teil enthält die Degradation von Transistorparametern. Die letzte Messung erfolgte nach SBD.

In der Abbildung 5.6 ist ein Beispiel für einen NMOS-Kurzkanaltransistor gezeigt, der einen erhöhten Off-Strom nach einem SBD im Gateoxid aufweist. Der Off-Strom ist definiert als der Drainstrom bei $V_d = V_{DD}$ und $V_g = 0 \text{ V}$. Im linken Teil der Abbildung 5.6 sind die Eingangskennlinien vor Beginn des Streßexperimentes und nach SBD gezeigt. Für $V_d = V_{DD} = 1,8 \text{ V}$ ist deutlich der erhöhte Drainstrom bei $V_g \leq 0 \text{ V}$ zu erkennen. Für $V_d = 0,1 \text{ V}$ ist dies nicht der Fall. Im rechten Teil der Abbildung ist die Degradation der Transistorparameter während des Stresses gezeigt. Die letzte Messung erfolgte, nachdem der SBD aufgetreten

war. Durch den SBD wurde der Off-Strom beeinflusst, nicht jedoch die Einsatzspannungen oder die Sättigungsdrainströme. Die sonstige Degradation der Transistorparameter ist durch den Streß selbst bedingt. Die Konfiguration des Streßexperimentes entspricht einem „Negative Bias Temperature Stress“ (NBTS) an einem NMOSFET.

Für eine genauere Analyse des GIDL bei NMOS-Feldeffekttransistoren kann die im vorhergehenden Kapitel beschriebene Meßroutine weiter modifiziert werden. Drain und Source werden zusammen an eine „Source Measurement Unit“ kontaktiert. Während der Unterbrechung des NBTS zur Bestimmung des Monitorstromes bei $V_g = -2$ V wird bei derselben Gatespannung der Strom an den Diffusionskontakten gemessen. Um die Potentialverhältnisse für GIDL zu erhalten, wird dabei an den Diffusionskontakten die Versorgungsspannung V_{DD} angelegt.

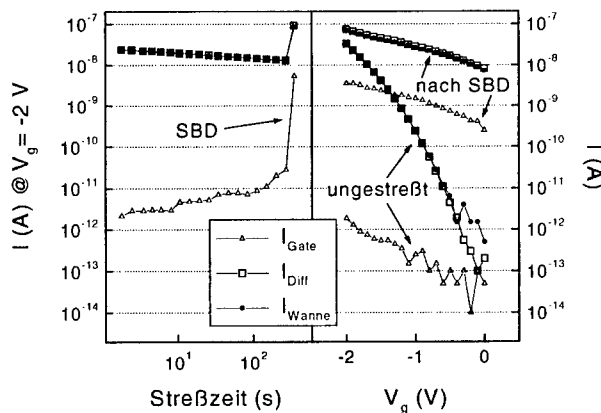


Abb. 5.7: „Gate-Induced Drain Leakage“ im direkten Vergleich mit einem SBD in einem NMOS-Kurzkanaltransistor. Im linken Teil der Abbildung sind die Ströme am Gate, an den Diffusionskontakten und am Wannenkontakt während des Stresses gezeigt. Im rechten Teil sind die Ströme im Gatespannungsbereich $-2 \leq V_g \leq 0$ dargestellt. „Gate-Induced Drain Leakage“ wird hier durch den SBD erhöht.

Das Ergebnis eines solchen Experimentes ist im linken Teil der Abbildung 5.7 gezeigt. Der Gateleckstrom I_{Gate} steigt in der Degradationsphase des SILC kontinuierlich an. In dieser Phase werden der Strom an den Diffusionskontakten I_{Diff} und der Strom am Wannenkontakt I_{Wanne} kleiner. Der SBD äußert sich durch den diskontinuierlichen Anstieg im Gateleckstrom. Im gezeigten Beispiel steigen die Ströme am Diffusionskontakt und am Wannenkontakt gleichzeitig sprunghaft an. Im rechten Teil der Abbildung 5.7 sind dieselben Ströme für einen Gatespannungsbereich von 0 V bis -2 V gezeigt, wobei an den Diffusionskontakten wieder

die Spannung V_{DD} anliegt. Es sind die Ströme vor Beginn des Stresses und nach dem SBD miteinander verglichen. Am ungestreßten Transistor wird ein normaler GIDL gemessen und der Gateleckstrom entspricht einem ca. 3 nm dicken Gateoxid. Nach dem SBD ist der Gateleckstrom dementsprechend angestiegen. Der GIDL ist ebenfalls sehr stark erhöht. Bei $V_g = 0$ V entspricht die Zunahme des GIDL sogar 5 Größenordnungen. Der Off-Strom des Transistors wird durch diesen stark erhöhten GIDL bestimmt.

5.2.4 Modell für den Einfluß des „Soft Breakdown“ auf „Gate-Induced Drain Leakage“

Im folgenden wird ein Modell dargelegt, das die Erhöhung des „Gate-Induced Drain Leakage“ durch einen „Soft Breakdown“ für NMOSFETs erklärt [5.12]. Der GIDL ist abhängig vom elektrischen Feld an der Substratoberfläche im Überlappbereich von Gate und Drain (siehe Kapitel 7.5 im Anhang A). Werden im Überlappbereich von Gate und Drain negative Oxidladungen generiert, dann erhöht sich der GIDL, da das elektrische Feld an der Substratoberfläche negativer wird. Ein entsprechender Effekt wird bei einem Hot-Carrier-Streß beobachtet. In das Oxid injizierte Elektronen, die in akzeptorartigen Haftstellen eingefangen werden, erhöhen in Drainnähe das vertikale elektrische Feld. Der erhöhte GIDL nach einem Hot-Carrier-Streß wurde sehr gut von A. v. Schwerin et al. beschrieben [5.28].

In der Abbildung 5.7 nimmt der GIDL in der Degradationsphase des SILC ab. Dies kann durch die Generation von positiven Ladungen im Oxid in der Nähe der Oxid/Substrat-Grenzfläche erklärt werden. Bei „Negative Bias Temperature Stress“ bildet das Substrat die Anode, und es ist bekannt, daß bei einem elektrischen Streß von SiO_2 positive Ladungen in der Nähe der Anode generiert werden [5.9, 5.29]. Die positiven Ladungen verkleinern das vertikale elektrische Feld an der Substratoberfläche und damit auch den GIDL.

Die Modellannahme ist nun, daß im Bereich des SBD-Pfades negative Ladungen auftreten. Die negativen Ladungen können, wie bereits oben erwähnt, Elektronen in akzeptorartigen Haftstellen sein. Akzeptorartige Haftstellen sind nach der Klassifizierung im Kapitel 2.1 neutrale Haftstellen für Elektronen. In den Kapiteln 5.1.1 und 5.1.3 wurde bereits erwähnt, daß neutrale Haftstellen während eines elektrischen Stresses im Oxid generiert werden. Deren Rolle in Bezug auf SILC, SBD oder auch HBD ist aber noch nicht gänzlich geklärt. Eine denkbare Quelle für negative Ladungen im Oxid bietet der SBD-Pfad selbst, der im allgemeinen als eine Kette von Haftstellen im Oxid verstanden wird, die eine Elektronenleitung von der Kathode zur Anode ermöglicht. Ist nun der SBD-Pfad durch das Oxid im Überlappbereich zwischen Gate und Drain lokalisiert, dann erhöhen die negativen Ladungen das vertikale elektrische Feld und den GIDL. Dies ist in der Abbildung 5.8 schematisch dargestellt. Zusätzlich sind die Verarmungszonen durch unterschiedliche Schattierungen skizziert. Zum besseren Verständnis des erhöhten GIDL ist auch die Bandverbiegung im n-dotierten Gebiet am Ort des SBD-Pfades angegeben.

In der Abbildung 5.9 sind Simulationen des GIDL für einen NMOSFET mit $W/L = 10/0,20 \mu\text{m}$ gezeigt. Der GIDL ohne Oxidladungen und der GIDL für Verteilungen an negativen Ladungen, die sich im Überlappbereich von Gate und Drain befinden, sind miteinander verglichen. Wie erwartet, steigt der GIDL mit zunehmender negativer

Oxidladung an. Für die laterale Ausdehnung des Überlappbereiches wurde $0,01\ \mu\text{m}$ gewählt. Die zweidimensionale Simulation zeigt den Strom über die gesamte Weite des Transistors von $10\ \mu\text{m}$. Für den GIDL ohne Ladungen ist dies korrekt. Der SBD tritt allerdings lokal auf. Damit trägt nur ein Teil der gesamten Transistorweite zum erhöhten GIDL bei.

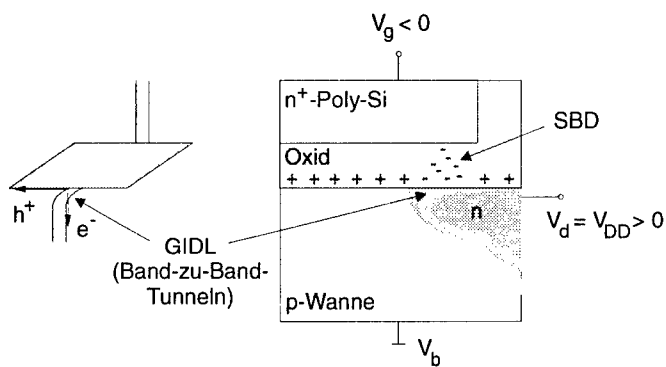


Abb. 5.8: Schematische Darstellung des Modells zur Erklärung des erhöhten „Gate-Induced Drain Leakage“ (GIDL) nach „Soft Breakdown“ (SBD) in einem NMOSFET.

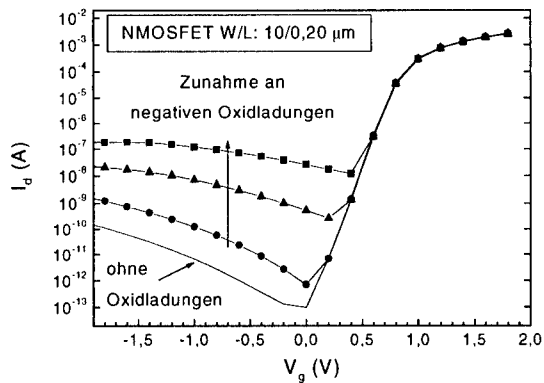


Abb. 5.9: Zweidimensionale Simulation des GIDL für einen NMOSFET, $W/L = 10/0,20\ \mu\text{m}$. Die negativen Oxidladungen sind im Bereich des Überlapp von Gate und Drain lokalisiert.

Für die Abschätzung der relevanten Fläche werden zwei Annahmen gemacht. Zum einen sei die Ausdehnung des SBD-Pfades symmetrisch in der Gateebene und zum anderen soll sich der SBD auf den Überlappbereich beschränken. Für die SBD-Fläche ergibt sich damit 10^{-12} cm^2 . Dies ist in Übereinstimmung mit Literaturwerten, die in Modellrechnungen ermittelt wurden [5.13, 5.15]. Der GIDL hat eine ähnliche Abhängigkeit vom elektrischen Feld wie der Fowler-Nordheim-Tunnelstrom. Konkret gilt [5.30]:

$$I_{\text{GIDL}} = W \cdot L_{\text{ov}} \int_0^{x_i} q \cdot a \cdot E^2(x) \cdot \exp\left[-\frac{b}{E(x)}\right] \cdot dx$$

$$a = \frac{2\pi \cdot q^2 \cdot \sqrt{m^*}}{9 \cdot h^2 \cdot E_{\text{gap}}^{1/2}}$$

$$b = \frac{\pi^2 \cdot E_{\text{gap}}^{3/2} \cdot \sqrt{m^*}}{h \cdot q}$$
(5.3)

$E(x)$ ist das elektrische Feld in der Tiefe x im Silizium und x_i ist die Tiefe, bei der die Ladungsträger vom Valenzband in das Leitungsband tunneln. W ist die Transistorweite und L_{ov} ist die laterale Länge des Überlappbereiches zwischen Gate und Drain. m^* ist die effektive Masse der Ladungsträger. Mit den obengenannten Annahmen beträgt die Ausdehnung des SBD-Pfades nur $0,01 \mu\text{m}$ von der gesamten Transistorweite von $10 \mu\text{m}$. Der tatsächlich durch SBD erhöhte GIDL muß somit um einen Faktor 10^3 größer sein. Eine einfache Abschätzung für die Oxidladung, welche notwendig ist, um einen Faktor 10^3 im GIDL zu erreichen, wird mit folgender Gleichung vorgenommen:

$$\frac{I_1}{I_2} = 10^3 = \frac{E_1^2 \cdot \exp(-b/E_1)}{E_2^2 \cdot \exp(-b/E_2)}$$
(5.4)

Wird für die effektive Elektronenmasse ein Wert zwischen $0,5 \cdot m_e$ und $1,0 \cdot m_e$ gewählt, dann ergibt sich mit $E_{\text{gap}} = 1,12 \text{ eV}$ für die Konstante b ein Wert zwischen 48 MV/cm und 67 MV/cm . In der Abbildung 5.10 ist für diese beiden Fälle der Verlauf von $E^2 \cdot \exp(-b/E)$ gegenüber dem elektrischen Feld E an der Siliziumoberfläche aufgetragen. Für den in der Abbildung 5.7 gezeigten NMOSFET ist bei $V_g = 0 \text{ V}$ und $V_d = 1,8 \text{ V}$ das elektrische Feld an der Substratoberfläche im Überlapp von Gate und Drain maximal 2 MV/cm . Ausgehend von diesem Wert ist in der Abbildung 5.10 die Zunahme des elektrischen Feldes gezeigt, die erforderlich ist, um eine Zunahme um 5 Größenordnungen (G.O.) zu erzielen. Dies ist exemplarisch für $m^* = 0,5 \cdot m_e$ dargestellt. Es ist außerdem die Zunahme des elektrischen Feldes eingetragen, die erforderlich ist, um lokal innerhalb des Überlapps von Gate und Drain von $0,01 \mu\text{m}$ denselben Effekt zu erzielen. Das elektrische Feld und damit die Dichte an negativen Oxidladungen muß also für den lokal auftretenden SBD um einen Faktor 2,8 größer sein, als wenn der GIDL über die gesamte Weite des Transistors generiert werden würde. Die Kurve für $m^* = m_e$ zeigt in der Abbildung 5.10 exemplarisch, daß für eine größere effektive Masse das erforderliche elektrische Feld kleiner wird. Zudem benötigt eine weitere Zunahme um 3 G.O. einen kleineren Faktor in der Dichte der Oxidladungen, da diese Zunahme in einem Kurvenbereich mit einer größeren Steigung erfolgt.

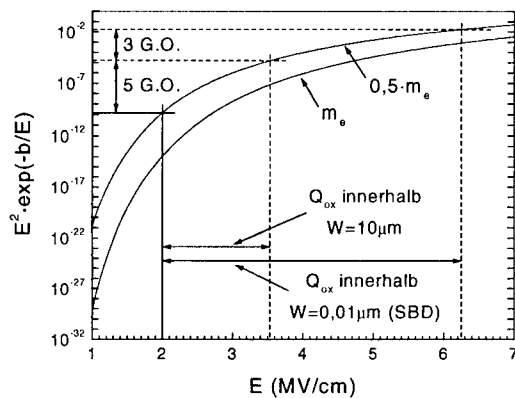


Abb. 5.10: Abschätzung für die erforderliche Erhöhung des elektrischen Feldes für eine Erhöhung des GIDL um 5 bzw. 8 Größenordnungen (G.O.). Die Abschätzung erfolgt anhand des feldabhängigen Anteils für GIDL in der Gleichung 5.3. Es sind exemplarisch die Verläufe für eine effektive Masse von $0,5 \cdot m_e$ und $1 \cdot m_e$ gezeigt. m_e ist die freie Elektronenmasse.

Eine vergleichbare Zunahme des GIDL, wie sie in der Abbildung 5.7 gezeigt ist, erfordert eine Dichte an negativen Oxidladungen von ca. $2 \cdot 10^{12} \text{ cm}^{-2}$ über die gesamte Weite des NMOSFETs. Dies zeigen die Simulationen. Die obige Abschätzung ergibt für die Lokalisierung im SBD-Bereich einen Faktor in der Größenordnung 1 und eine Dichte an negativen Oxidladungen in der Größenordnung 10^{12} cm^{-2} . Die Abschätzung zeigt also, daß der Effekt mit einer Anreicherung von negativen Ladungen im SBD-Pfad oder in der Nähe des SBD-Pfades erklärt werden kann. Der Grund ist die starke Abhängigkeit des GIDL vom elektrischen Feld, die derjenigen für die Fowler-Nordheim-Tunnelstromdichte entspricht und im Bereich kleiner elektrischer Felder einen hohen Gradienten aufweist.

5.2.5 „Soft Breakdown“ in unterschiedlichen Bereichen des Feldeffekttransistors

Aufgrund des lokalen Charakters des „Soft Breakdown“ kann der SBD-Pfad im Oxid in unterschiedlichen Bereichen des Transistorkanals auftreten. In der Abbildung 5.11 sind die $I(V_g)$ -Kurven eines NMOS-Kurzkanaltransistors nach SBD für den Gateleckstrom, den Drainstrom, den Sourcestrom und den Strom am Wannenkontakt gezeigt. Im rechten Teil ist der NMOSFET in Vorwärtsrichtung geschaltet, d.h. Source ist geerdet und am Drain liegt die Versorgungsspannung V_{DD} an. Im linken Teil ist der NMOSFET in Rückwärtsrichtung geschaltet, d.h. der Drainkontakt ist geerdet und am Sourcekontakt liegt die

Versorgungsspannung V_{DD} an. Der erhöhte GIDL ist nur in Vorwärtsrichtung meßbar. Der SBD-Pfad ist also im Überlappbereich von Gate und Drain lokalisiert. Im rechten Teil der Abbildung 5.11 ist erkennbar, daß der Gateleckstrom bei $V_g = 0$ V bereits sehr hoch ist. Dies zeigt auch, daß ein Leckstrompfad zwischen dem Gate und dem Drainkontakt gebildet wurde und nicht zwischen dem Gate und der geerdeten Wanne.

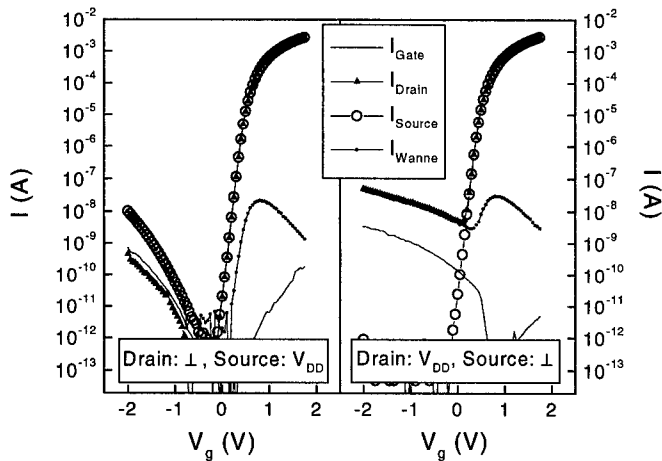


Abb. 5.11: Rechts: Ströme am Gate, an Drain, an Source und am Wannenkontakt für einen NMOSFET mit SBD im Überlapp von Gate und Drain. Links: Ströme für denselben NMOSFET, der nun in Rückwärtsrichtung gepolt ist. An Source ist kein erhöhter GIDL meßbar.

In der Abbildung 5.12 sind entsprechende Messungen für einen anderen NMOS-Kurzkanaltransistor nach SBD gezeigt. Hier wird in Rückwärtsrichtung ein erhöhter GIDL gemessen, d.h. der SBD-Pfad ist in diesem Fall im Überlappbereich von Gate und Source lokalisiert. Bei Drain tritt keine Degradation auf. Dieses Beispiel zeigt auch, daß je nach Oxidladungsdichte der GIDL bei $V_g = 0$ V auch vergleichbar zum Gateleckstrom sein kann. Es wurden auch Transistoren gefunden, für die der Gateleckstrom größer als der erhöhte GIDL ist. In einem solchen Fall wird der Off-Strom durch den Gateleckstrom bestimmt. In der Abbildung 5.13 ist ein NMOS-Kurzkanaltransistor gezeigt, für den der SBD-Pfad zwischen den p-n-Übergängen von Drain und Source lokalisiert ist. Es wird weder in Vorwärtsrichtung noch in Rückwärtsrichtung ein erhöhter GIDL gemessen. In den beiden Betriebsarten verhält sich jedoch der Gateleckstrom ähnlich wie bei Messungen an MOS-Kapazitäten. Dies bedeutet, daß der Leckstrompfad tatsächlich zwischen Gate und Wanne verläuft.

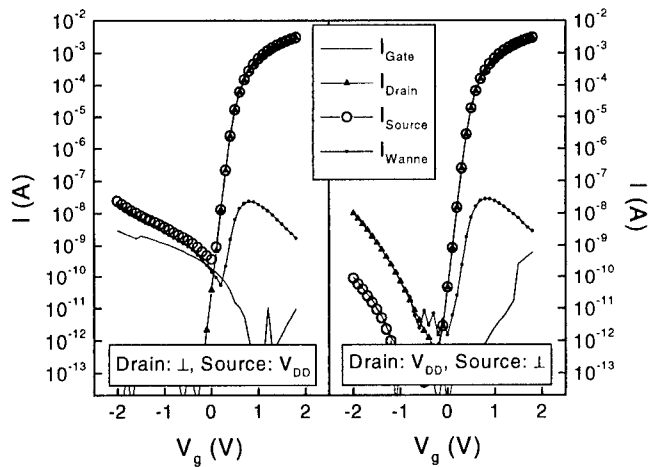


Abb. 5.12: Entsprechend zur Abbildung 5.11 sind die Ströme eines anderen NMOSFETs dargestellt. Der SBD-Pfad im Oxid ist hier im Überlappbereich von Gate und Source lokalisiert

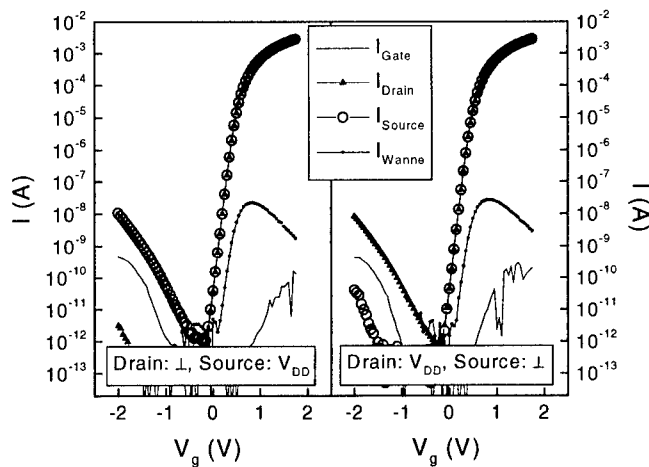


Abb. 5.13: Entsprechend zur Abbildung 5.11 sind die Ströme eines anderen NMOSFETs dargestellt. Der SBD-Pfad im Oxid ist zwischen den p-n-Übergängen von Drain und Source lokalisiert.

5.2.6 Einfluß auf die Funktionalität eines NMOSFETs

Mit Ausnahme des Off-Stromes weisen alle gemessenen Transistorparameter keine Degradation durch den SBD auf [5.12]. Nach dem Modell, das im Kapitel 5.2.4 vorgestellt wurde, tritt die Erhöhung des GIDL nur auf, wenn der SBD-Pfad im Gateoxid im Überlappbereich von Gate und Drain lokalisiert ist. Bei entsprechenden Streßexperimenten an NMOS-Langkanaltransistoren mit $L = 10\ \mu\text{m}$ wurde keine Degradation der überprüften Transistorparameter, auch nicht des Off-Stromes, gefunden [5.12]. Der Grund ist, daß mit zunehmender Gatelänge die Wahrscheinlichkeit abnimmt, daß der SBD-Pfad im Überlappbereich von Gate und Drain lokalisiert ist.

5.2.7 Einfluß auf die Funktionalität eines PMOSFETs

Bei entsprechenden Streßexperimenten an PMOS-Feldeffekttransistoren wurde kein erhöhter GIDL gefunden [5.19]. Mit dem im Kapitel 5.2.4 eingeführten Modell ist erklärbar, daß bei einem PMOSFET ein solcher Effekt nicht auftreten kann. Die positiven Ladungen, die während der SILC-Phase erzeugt werden, führen beim PMOSFET zu einem leichten Anstieg des GIDL, da hier die Potentialverhältnisse umgekehrt zum NMOSFET sind. Dies wird auch während des elektrischen Stresses beobachtet [5.19]. Dementsprechend führen negative Ladungen im Bereich des SBD-Pfades zu einer Verkleinerung des GIDL. Der Off-Strom des PMOSFETs bleibt damit durch den regulären Drainstrom bestimmt und ändert sich nicht. Vergleichbar dem NMOSFET wurde eine Degradation anderer Transistorparameter durch SBD nicht beobachtet [5.19]. Somit ist festzustellen, daß der SBD keinen Einfluß auf die Funktionalität eines PMOSFETs hat.

5.3 Modell des selbstlimitierenden „Soft Breakdown“-Pfades

Bei der Beschreibung der Charakteristika des „Soft Breakdown“ wurde bereits erwähnt, daß der Gateleckstrom nach SBD starke Fluktuationen aufweist. Unter bestimmten Meßbedingungen wurden auch Fluktuationen zwischen zwei diskreten Niveaus gemessen [5.13, 5.15, 5.16]. Brière et al. führen die Fluktuationen im Gateleckstrom nach SBD auf die Änderung des Ladungszustandes einer Haftstelle zurück [5.15]. Das im Kapitel 5.2.4 eingeführte Modell zur Erklärung des erhöhten GIDL bei NMOSFETs geht von negativen Ladungen im Bereich des SBD-Pfades aus. Damit wird jedoch an dieser Stelle das elektrische Feld zwischen Kathode und dem Ort der negativen Ladungen erniedrigt, was zu einem reduzierten Fluß an Elektronen führen muß. Wenn die Elektronen akzeptorartige Haftstellen besetzen, die selbst ein Bestandteil des leitfähigen Pfades zwischen Kathode und Anode sind, dann wirkt die Coulombabschirmung dem Elektronentransport im Pfad entgegen. Der SBD-Pfad ist damit selbstlimitierend. Dieses Modell des selbstlimitierenden SBD-Pfades wurde in [5.19] erstmals diskutiert und anhand des GIDL eines NMOSFETs überprüft.

In der Abbildung 5.14 ist das Resultat eines entsprechenden Streßexperimentes an einem NMOSFET gezeigt, das im Kapitel 5.2.3 diskutiert wurde. Die graphische Darstellung ist identisch mit der Abbildung 5.7. Im Beispiel in der Abbildung 5.14 traten während der

Messung der $I_g(V_g)$ -Kurve Fluktuationen im Gateleckstrom auf. Dieser nimmt bei ca. -0,7 V ab und bei ca. -1,2 V wieder leicht zu. Der Strom am Wannenkontakt verhält sich genau umgekehrt. Die Änderung des Stromes an den Diffusionskontakten ist nicht so ausgeprägt, verhält sich aber entsprechend wie der Strom am Wannenkontakt. Das ist darauf zurückzuführen, daß bis zum ersten Sprung im Gateleckstrom dieser in der gleichen Größenordnung wie der GIDL liegt. Damit wird der Strom an den Diffusionskontakten in diesem Spannungsbereich vom Gateleckstrom stark beeinflusst. Eine Abnahme des Gateleckstromes und eine Zunahme des GIDL kompensieren eine Änderung des Stromes an den Diffusionskontakten nahezu. Die Messung zeigt, daß eine Abnahme des Gateleckstromes mit einer gleichzeitigen Zunahme des GIDL verbunden ist. Dies bestätigt das Modell des selbstlimitierenden SBD-Pfades. Eine oder mehrere zusätzliche negative Ladungen reduzieren den Elektronentransport durch den SBD-Pfad, erhöhen aber das elektrische Feld und damit den GIDL.

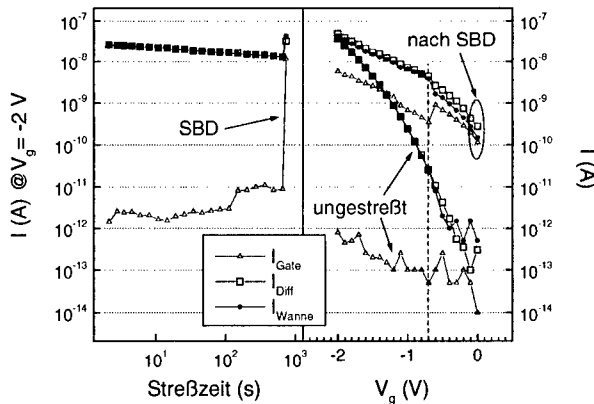


Abb. 5.14: Streßexperiment an einem NMOS-Kurzkanaltransistor, das dem in der Abbildung 5.7 entspricht. In der $I_g(V_g)$ -Kurve nach SBD treten hier Fluktuationen auf. Der GIDL verhält sich umgekehrt wie der Gateleckstrom.

Aus den bisher gewonnenen Ergebnissen und den Erkenntnissen in der Literatur können folgende Charakteristika für den SBD-Pfad zusammengefaßt werden. Während des elektrischen Stresses werden im Oxid akzeptorartige und donatorartige Haftstellen generiert. Die akzeptorartigen Haftstellen, also neutrale Haftstellen für Elektronen, ermöglichen einen Elektronentransport zwischen den Elektroden. Durch den SBD-Pfad findet ein gegenüber dem „Stress-Induced Leakage Current“ (SILC) erhöhter Transport von Elektronen statt. Im SBD-Pfad oder in dessen Nähe sind die meisten akzeptorartigen Haftstellen mit Elektronen besetzt. Aufgrund des repulsiven Coulombpotentials der Elektronen und der Verkleinerung des

elektrischen Feldes ist der SBD-Pfad selbstlimitierend. Der erhöhte Transport an Elektronen ist damit sehr sensitiv gegenüber dem Besetzungszustand der Haftstellen. Dies führt zu den Fluktuationen im Gateleckstrom. Im Kapitel 4.4 wurde gezeigt, daß der Unterschied zwischen dem Transport von Elektronen bei SILC und nach SBD mit der Ionisationsenergie der Haftstellen in Verbindung gebracht werden kann.

5.4 Beschleunigung der Zeit bis zum „Soft Breakdown“ im Vergleich mit dem dielektrischen Durchbruch

Im folgenden werden die Temperaturbeschleunigung und die Feldbeschleunigung des „Soft Breakdown“ (SBD) und des dielektrischen Durchbruches (HBD) getrennt voneinander untersucht. Viele Autoren definieren die Zeit bis zum Durchbruch in dünnen Oxiden als die Zeit bis zum ersten Durchbruchereignis, SBD oder HBD [5.31, 5.32, 5.33]. Insbesondere bei Streßexperimenten mit hohen elektrischen Feldern kann ein dielektrischer Durchbruch auftreten, ohne daß zuvor ein SBD-Ereignis gemessen wurde.

Es wurden NMOS-Kapazitäten in Akkumulation gestreßt, d.h. die Polarität der Gatespannung war negativ. Dies ist auch der relevante Streßmodus für den durch SBD erhöhten GIDL in NMOSFETs. Wie im Kapitel 5.2 gezeigt wurde, degradiert der SBD den NMOSFET nur, wenn der SBD-Pfad im Überlappbereich von Gate und Drain lokalisiert ist. Eine Potentialdifferenz zwischen dem Gate und dem Drain liegt im Off-Zustand vor. Für einen NMOSFET ist diese negativ. Außerdem wurde im Kapitel 4.7 gezeigt, daß für die Untersuchung der Degradationsbeschleunigung des Oxides ein Streß in Akkumulation zu bevorzugen ist.

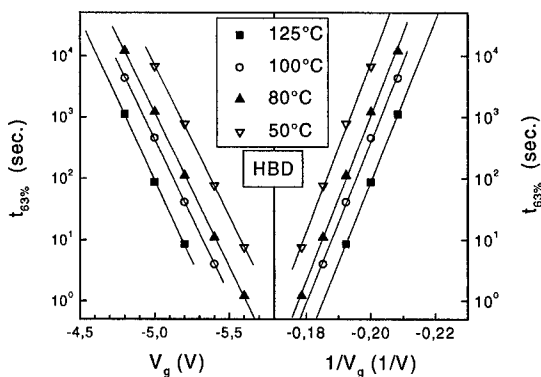


Abb. 5.15: $t_{63\%}$ bis zum „Hard Breakdown“ (HBD) für verschiedene Temperaturen und in Abhängigkeit von der Gatespannung V_g (links) sowie in Abhängigkeit von $1/V_g$ (rechts). Die Oxiddicke ist 3,4 nm.

5.4.1 Beschleunigung in Abhängigkeit vom elektrischen Feld

Die Streßparameter sind so gewählt, daß Durchbruchzeiten zwischen einer Sekunde und 10^4 Sekunden gemessen wurden. Dies entspricht dem experimentellen Fenster, das Streßexperimenten an Wafern („wafer level test“) zur Verfügung steht. Längere Zeiten sind nur mit Langzeitexperimenten und gehäuteten Testmodulen („package level test“) sinnvoll. Als charakteristische Zeit für den intrinsischen Durchbruch wurde jeweils die Zeit gewählt, bei der für 63% der getesteten Kapazitäten das jeweilige Durchbruchereignis meßbar war. In der Abbildung 5.15 ist gezeigt, daß die gemessenen Zeiten bis zum HBD in einer halblogarithmischen Graphik sowohl gegenüber dem elektrischen Feld E als auch gegenüber $1/E$ sehr gut mit Geraden angepaßt werden können [5.34]. Dieses Problem wurde bereits im Kapitel 2.3.2 diskutiert. Eine Überprüfung des E-Modells und des $1/E$ -Modells ist nur mit Langzeitexperimenten möglich. Gleiches gilt für den SBD, wie der Abbildung 5.16 zu entnehmen ist. Im folgenden wird das E-Modell verwendet, um das unterschiedliche Beschleunigungsverhalten zu charakterisieren. Die Unterschiede zwischen SBD und HBD sind jedoch Meßergebnisse und daher modellunabhängig.

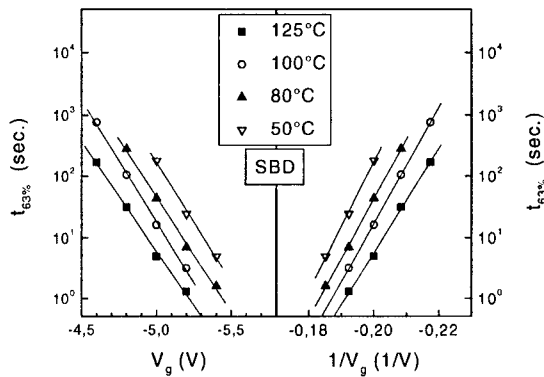


Abb. 5.16: $t_{63\%}$ bis zum „Soft Breakdown“ (SBD) für verschiedene Temperaturen und in Abhängigkeit von der Gatespannung V_g (links) sowie in Abhängigkeit von $1/V_g$ (rechts). Die Oxiddicke ist 3,4 nm.

In der Abbildung 5.17 sind die Feldbeschleunigungsfaktoren γ des E-Modells gezeigt. Die Beschleunigungsfaktoren wurden aus den Steigungen in den Abbildungen 5.15 und 5.16 ermittelt. Die Fehlerbalken entsprechen der Genauigkeit der Geraden durch die Meßpunkte. Der SBD weist im gesamten untersuchten Temperaturbereich eine deutlich kleinere Beschleunigung durch das elektrische Feld auf als der HBD. Es war bisher nicht bekannt, daß ein deutlicher Unterschied im Beschleunigungsverhalten auftreten kann und wurde in [5.34]

erstmals gezeigt. Damit steht fest, daß SBD- und HBD-Ereignisse getrennt voneinander charakterisiert werden müssen.

Nach dem thermochemischen E-Modell folgt der Feldbeschleunigungsfaktor γ in einer Arrhenius-Graphik einer ansteigenden Geraden [5.35]. Mit ansteigender Temperatur nimmt also die Feldbeschleunigung im E-Modell ab. Kimura hat dies am Beispiel eines 9,9 nm dicken Oxides bestätigt [5.36]. Suehle et al. fanden für ein 15 nm dickes und ein 22,5 nm dickes Oxid keine Temperaturabhängigkeit [5.37]. Dieselben Autoren bestätigten dies in einer weiteren Arbeit für Oxiddicken von 6,5 nm, 9 nm, 15 nm und 22 nm [5.38]. Yassine et al. bestimmten eine lineare Zunahme des Feldbeschleunigungsfaktors γ in einer Arrhenius-Graphik für ein 3,9 nm dickes Oxid [5.39] und Bruyère et al. fanden hingegen eine Abnahme des Feldbeschleunigungsfaktors γ in einer Arrhenius-Graphik für ein 3,5 nm dickes Oxid [5.40]. Die Temperaturabhängigkeit des Feldbeschleunigungsfaktors für den HBD entspricht bei Bruyère et al. dem in der Abbildung 5.17 gezeigten Verhalten für den HBD.

Das thermochemische E-Modell kann erklären, daß der Feldbeschleunigungsfaktor in einem begrenzten Temperaturbereich unabhängig von der Temperatur ist. Dafür muß eine Mischung aus mehreren Bindungszuständen zugrunde gelegt werden, die jeweils unter dem Einfluß des elektrischen Feldes aufbrechen und zu einem dielektrischen Durchbruch führen („mixing of disturbed bonding states“) [5.35]. Das ursprüngliche thermochemische E-Modell bezieht sich demgegenüber nur auf Si-Si-Bindungen. Bruyère et al. zeigten, daß mit einer geeigneten Variation von Parametern des E-Modells mit mehreren Bindungszuständen sogar die beobachtete Abnahme des Feldbeschleunigungsfaktors γ in einer Arrhenius-Graphik erklärt werden kann [5.40]. Das Temperaturverhalten des Feldbeschleunigungsfaktors für den HBD ist somit qualitativ erklärbar.

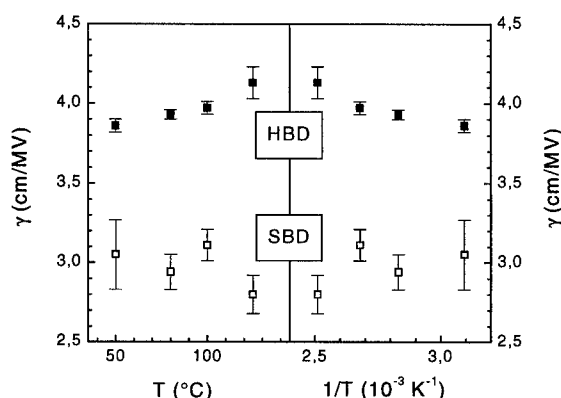


Abb. 5.17: Feldbeschleunigungsfaktoren γ für SBD und HBD in Abhängigkeit von der Temperatur T und von $1/T$. Die Oxiddicke ist 3,4 nm.

In der vorliegenden Arbeit wurde für die Untersuchung des Beschleunigungsverhaltens ein in N_2O -Atmosphäre nitriertes Oxid verwendet. Vor der Nitridierung erfolgte eine Trockenoxidation und eine Naßoxidation des Siliziumsubstrates. Die Nitridierung könnte ein Grund dafür sein, daß eine Mischung aus mehreren Bindungszuständen für den dielektrischen Durchbruch relevant ist und der Feldbeschleunigungsfaktor ein Temperaturverhalten zeigt, das bisher nur mit dem E-Modell für eine Mischung aus mehreren Bindungszuständen („mixing of disturbed bonding states“) erklärt werden kann.

5.4.2 Beschleunigung in Abhängigkeit von der Temperatur

In diesem Kapitel wird die Beschleunigung der elektrischen Degradation durch die Temperatur betrachtet. Im linken Teil der Abbildung 5.18 sind für die verschiedenen Streßspannungen die Zeiten $t_{63\%}$ bis zum HBD halblogarithmisch gegenüber $1/T$ aufgetragen. Zwischen 50°C und 150°C lassen sich die Daten sehr gut mit einer Geraden anpassen. Die Datenpunkte für $T = 30^\circ\text{C}$ weichen davon jedoch ab. Im rechten Teil der Abbildung 5.18 sind dieselben Zeiten halblogarithmisch gegenüber der Temperatur dargestellt. Die Zeit bis zum HBD weist demnach eher eine lineare Abhängigkeit von der Temperatur auf.

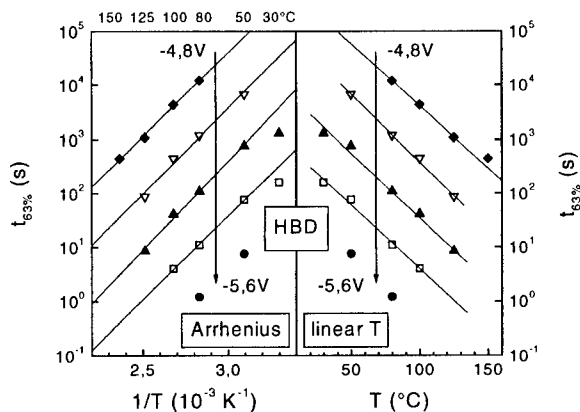


Abb. 5.18: $t_{63\%}$ bis zum „Hard Breakdown“ (HBD) in Abhängigkeit von der Temperatur und für verschiedene Streßspannungen. Links: Arrhenius-Graphik. Rechts: Lineare Darstellung gegenüber der Temperatur T . Die Oxiddicke ist 3,4 nm.

Das E-Modell berücksichtigt die Temperaturbeschleunigung gemäß einer Arrhenius-Gleichung (siehe Kapitel 2.3.2). In der Literatur wurde bereits für dünne Oxide eine Abweichung der Degradation vom Arrhenius-Verhalten berichtet [5.41]. Kaczer et al.

beobachteten für einen weiten Oxiddickenbereich von 2,2 nm bis 13,8 nm eine lineare Abhängigkeit der Zeit bis zum Durchbruch von der Temperatur [5,42]. Die Autoren führen dies darauf zurück, daß sowohl die Generationsrate von Haftstellen als auch die kritische Haftstellendichte für HBD temperaturabhängig sind. Es existiert jedoch kein Modell, das eine Abweichung vom Arrhenius-Verhalten erklärt.

Für den SBD sind die Zeiten $t_{63\%}$ in der Abbildung 5.19 halblogarithmisch gegenüber $1/T$ und gegenüber T aufgetragen. In beiden graphischen Darstellungen können die Meßdaten sehr gut mit Geraden angepaßt werden. Damit bleibt auch für den SBD ungeklärt, ob die Beschleunigung durch die Temperatur mit einer Arrhenius-Gleichung korrekt beschrieben wird.

Für die Zuverlässigkeitsuntersuchungen von Oxiden ist der Temperaturbereich zwischen ca. 50°C und 150°C interessant, da sich die Betriebstemperatur und die Temperatur im Streßexperiment in der Regel in diesem Bereich befinden. Im Bereich von 50°C bis 150°C lassen sich die Meßdaten in der Arrhenius-Graphik sowohl für den SBD als auch für den HBD gut mit Geraden beschreiben. Damit kann in diesem Temperaturbereich eine Aktivierungsenergie E_a bestimmt werden. Dies ist zulässig, obwohl die obige Diskussion zeigt, daß eine Arrhenius-Gleichung möglicherweise keine allgemeingültige Beschreibung der Beschleunigung der Oxiddegradation in Abhängigkeit von der Temperatur ergibt.

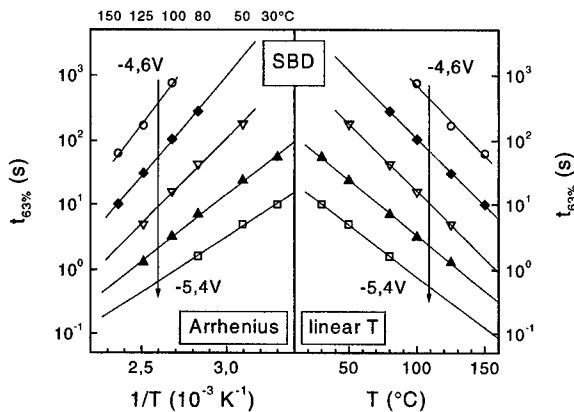


Abb. 5.19: $t_{63\%}$ bis zum „Soft Breakdown“ (SBD) in Abhängigkeit von der Temperatur und für verschiedene Streßspannungen. Links: Arrhenius-Graphik. Rechts: Lineare Darstellung gegenüber der Temperatur T . Die Oxiddicke ist 3,4 nm.

In der Abbildung 5.20 sind die Aktivierungsenergien in Abhängigkeit vom elektrischen Feld im Oxid dargestellt. Die Fehlerbalken entsprechen wieder der Genauigkeit der Geraden durch die Meßpunkte. Der HBD weist praktisch eine konstante Aktivierungsenergie von ca. 0,27 eV auf. Ähnlich wie für die Temperaturabhängigkeit des Feldbeschleunigungsfaktors γ gibt es in der Literatur unterschiedliche Aussagen über die Abhängigkeit der Aktivierungsenergie vom elektrischen Feld im Oxid. Das E-Modell sagt eine lineare Abnahme der Aktivierungsenergie mit dem elektrischen Feld voraus [5.35]. Kimura hat das Modell für ein 9,9 nm dickes Oxid experimentell bestätigt und aus den Daten folgende Gleichung für den HBD bestimmt [5.36]:

$$E_a(\text{eV}) \approx 1,15 - 0,07 \cdot E_{\text{ox}}(\text{MV/cm}) \quad (5.5)$$

Yassine et al. beobachteten für ein 3,9 nm dickes Oxid ebenfalls eine Abnahme der Aktivierungsenergie des HBD mit dem elektrischen Feld [5.39]. Eine mit dem elektrischen Feld abnehmende Aktivierungsenergie für den HBD wurde auch von Bruyère et al. für ein 3,5 nm dickes Oxid gemessen [5.40]. In der Abbildung 5.20 sind die gemessenen Aktivierungsenergien mit den experimentellen Ergebnissen der oben erwähnten Autoren verglichen.

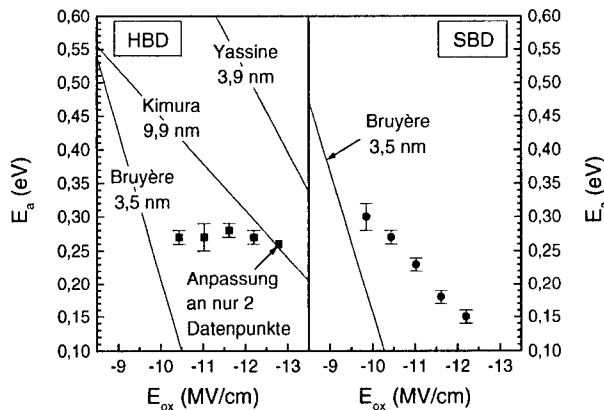


Abb. 5.20: Aktivierungsenergien für HBD und SBD in Abhängigkeit vom elektrischen Feld im Oxid. Die Oxiddicke ist 3,4 nm. Die Daten sind mit experimentellen Ergebnissen anderer Autoren [5.36, 5.39, 5.40] verglichen.

Eine konstante Aktivierungsenergie, wie sie im Fall des HBD gemessen wurde, kann in einem begrenzten Feldstärkebereich mit dem E-Modell für eine Mischung aus mehreren Bindungszuständen („mixing of disturbed bonding states“) erklärt werden [5.35]. Dieses Modell wurde bereits im vorhergehenden Kapitel zur qualitativen Erklärung des ungewöhnlichen Temperaturverhaltens der Feldbeschleunigung herangezogen. Damit ergibt

sich eine konsistente Erklärung für die Feldabhängigkeit der Aktivierungsenergie und die Temperaturabhängigkeit der Feldbeschleunigung, welche die experimentellen Daten für den HBD in den Abbildungen 5.17 und 5.20 aufweisen.

Die experimentellen Ergebnisse für den SBD werden im folgenden Kapitel diskutiert.

5.4.3 Analyse des „Soft Breakdown“

In der Abbildung 5.21 sind die Aktivierungsenergien für den „Soft Breakdown“ (SBD) gegenüber der Temperatur aufgetragen und durch eine Gerade angepaßt. Die angepaßte Gerade beschreibt folgende Gleichung:

$$E_a(\text{eV}) \approx 0,98 - 0,068 \cdot E_{\text{ox}}(\text{MV/cm}) \quad (5.6)$$

Als Vergleich sind in der Abbildung 5.21 auch die experimentellen Daten von Kimura [5.36] für den „Hard Breakdown“ (HBD) aufgeführt. Diese Daten ergeben eine Gerade gemäß Gleichung 5.5 und stimmen mit dem thermochemischen E-Modell überein, das besagt, daß die Aktivierungsenergie linear mit dem elektrischen Feld abnimmt. In diesem Modell ist die Steigung der Geraden durch das Dipolmoment des Bindungskomplexes O_3Si bestimmt [5.35]. Der Ordinatenabschnitt beträgt 1,15 eV und ist durch die Bindungsenergie der Si-Si-Bindung von 450 kJ/mol vorgegeben [5.35]. Ein Viertel pro Mol entfällt auf eine Si-Si-Bindung in Siliziumdioxid, das aus SiO_4 -Tetraedern besteht. Dies ist ca. 1,15 eV [5.35]. Die entsprechende Berechnung ergibt für den Wert 0,98 eV in der Gleichung 5.6 eine Bindungsenergie von ca. 378 kJ/mol.

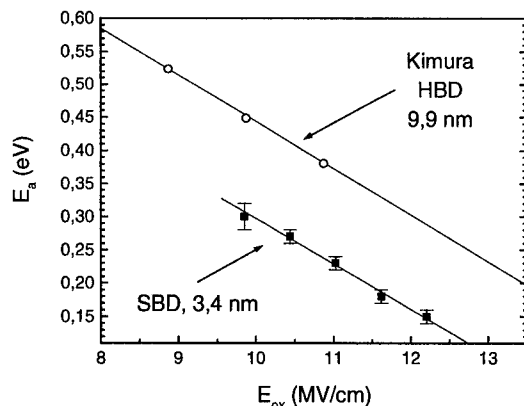


Abb. 5.21: Aktivierungsenergie für den SBD im Vergleich zu den experimentellen Daten von Kimura für einen HBD [5.36], die das E-Modell bestätigen.

Der SBD kann demnach in Analogie zum thermochemischen E-Modell mit dem Aufbrechen einer schwächeren Bindung in Verbindung gebracht werden. Es wird gezeigt, daß dies H-Si-Bindungen und H-O-Bindungen sein können.

Modellvorstellung für die Ausbildung eines SBD-Pfades

Im folgenden wird anhand der experimentellen Ergebnisse eine Modellvorstellung für die Ausbildung eines SBD-Pfades entwickelt [5.34]. Der Ladungsträgertransport kann aufgrund von Fragmenten von H-Si-Bindungen und H-O-Bindungen erfolgen. Ein Wasserstoffatom bildet dabei eine donatorartige Haftstelle für Elektronen ($X^+ + e^- \leftrightarrow X^0$). Das gebundene Siliziumatom, z.B. in einem ($O_3 \equiv Si \cdot$)-Komplex, und das gebundene Sauerstoffatom, z.B. in einem ($O_3 \equiv Si-O \cdot$)-Komplex, bilden jeweils akzeptorartige Haftstellen für Elektronen ($X^0 + e^- \leftrightarrow X^-$). Ein Ladungstransfer zwischen diesen Haftstellen ist möglich, wenn ein bestimmter Abstand zwischen den Haftstellen unterschritten wird. Das in der Literatur anerkannte „Percolation Model“ [5.23, 5.24] basiert auf der Existenz eines solchen minimalen Abstandes. Eine Anreicherung von negativen Ladungen im SBD-Pfad ist möglich, wenn die Reaktionsrate der Emission von Elektronen aus einer besetzten Haftstelle Si^- oder O^- kleiner ist als die Reaktionsrate der Besetzung einer leeren Haftstelle H^+ , $Si \cdot$ oder $O \cdot$. Die Anreicherung von negativen Ladungen im SBD-Pfad wurde im Kapitel 5.2.4 zur Erklärung des erhöhten „Gate-Induced Drain Leakage“ (GIDL) angenommen und im Kapitel 5.3 mit experimentellen Ergebnissen untermauert.

Die obige Modellvorstellung bedingt, daß die Ausbildung des SBD-Pfades durch die Diffusion von Wasserstoff in Siliziumdioxid beeinflusst werden kann. Die gemessenen Aktivierungsenergien für den SBD wurde in der Abbildung 5.21 mit einer Geraden angepaßt. Im Gegensatz dazu zeigt die Abbildung 5.22 eine Anpassung für eine Boltzmann-Gleichung zwischen zwei Grenzwerten. Die Gleichung lautet:

$$E_a(V_g) = \frac{E_{\max} - E_{\min}}{1 + \exp\left[\frac{V_g - V_0}{dV_g}\right]} + E_{\min} \quad (5.7)$$

Die Formparameter sind $V_0 = 5,03 \pm 0,03$ eV und $dV_g = 0,20 \pm 0,04$ eV. Die aus der Anpassung gewonnenen Grenzen für die Aktivierungsenergie sind $E_{\min} = 0,12 \pm 0,02$ eV und $E_{\max} = 0,32 \pm 0,01$ eV. Die Aktivierungsenergie für H_2 -Diffusion in SiO_2 beträgt 0,446 eV [5.43, 5.44]. Für die H-Diffusion in SiO_2 beträgt die Aktivierungsenergie 0,18 eV [5.44]. Gemäß der obigen Modellvorstellung ist der obere Grenzwert durch die Diffusion von H_2 und atomarem Wasserstoff erklärbar. Die lineare Abnahme der Aktivierungsenergie ist bei höheren Feldern analog zum thermochemischen E-Modell. Der untere Grenzwert ist dann durch die Diffusion von atomarem Wasserstoff gegeben, wenn bei hohen elektrischen Feldern die meisten H-Si-Bindungen und H-O-Bindungen aufgebrochen sind.

Die Anpassung der Daten in der Abbildung 5.22 ergibt, daß im Gegensatz zur Abbildung 5.21 eine Tangente durch den Wendepunkt zu verwenden ist. Diese Tangente ist mit folgender Gleichung für die feldabhängige Aktivierungsenergie verbunden:

$$E_a(\text{eV}) \approx 1,08 - 0,078 \cdot E_{\text{ox}}(\text{MV/cm}) \quad (5.8)$$

Die geringe Zunahme der Steigung erhöht die Aktivierungsenergie auf 1,08 eV. Dies entspricht einer Bindungsenergie von 417 kJ/mol. Die Bindungsenergie der H-Si-Bindung beträgt 300 kJ/mol und die Bindungsenergie der H-O-Bindung ist 427,5 kJ/mol [5.45]. Die gemäß Gleichung 5.6 und Gleichung 5.8 bestimmten Bindungsenergien liegen zwischen diesen beiden Werten. Damit ist eine Beteiligung beider Bindungen, H-Si und H-O, wahrscheinlich.

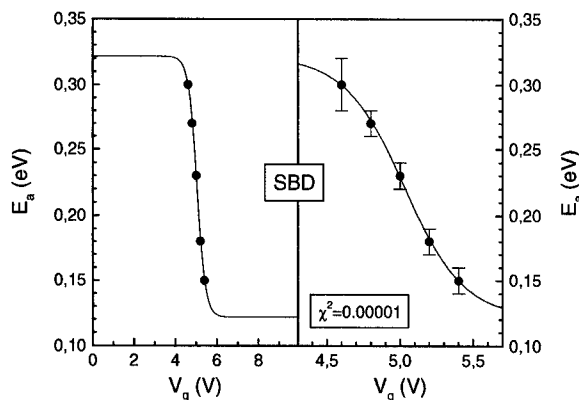


Abb. 5.22: Anpassung der spannungsabhängigen Aktivierungsenergie des „Soft Breakdown“ (SBD) mit einer Boltzmann-Gleichung.

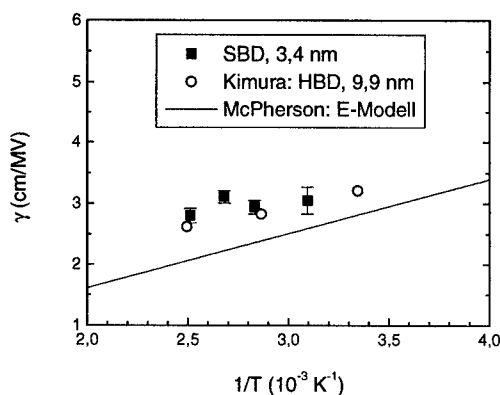


Abb. 5.23: Arrhenius-Graphik für die Feldbeschleunigung des SBD. Die Daten sind mit den experimentellen Daten von Kimura für einen HBD [5.36] und dem thermochemischen E-Modell [5.35] verglichen.

Die Daten der Feldbeschleunigung für den SBD weisen eine gute Übereinstimmung mit den Daten von Kimura [5.36] auf, die für einen HBD und ein 9,9 nm dickes Oxid bestimmt wurden. Dem Vergleich in der Abbildung 5.21 ist die Vorhersage des thermochemischen E-Modells [5.35] hinzugefügt. Die Steigung der Geraden wird in diesem Modell wiederum durch das Dipolmoment des $(\text{O}_3\equiv\text{Si})$ -Komplexes bestimmt [5.35].

Die Ursache für die Übereinstimmung der Feldbeschleunigung für SBD und HBD und deren geringfügige Abweichung vom Modell in der Abbildung 5.23 ist noch nicht bekannt. Eine qualitative Erklärung kann gegeben werden, wenn zusätzlich zur obigen Modellvorstellung für den SBD die folgenden beiden Annahmen gemacht werden.

- 1) Das Aufbrechen einer H-Si-Bindung oder einer H-O-Bindung führt nicht unmittelbar zu einem HBD, kann jedoch das Aufbrechen einer Si-Si-Bindung unterstützen, was nach dem E-Modell die Ursache für den HBD ist.
- 2) Die Feldbeschleunigung für das Aufbrechen der H-X-Bindungen ist im Vergleich zur Si-Si-Bindung höher. Das X steht stellvertretend für Silizium (Si) oder Sauerstoff (O).

Die erste Annahme ist sinnvoll, da der Wasserstoff in jedem Fall nur eine freie Bindung absättigt. Ein Aufbrechen einer solchen Bindung führt nicht zur Zerstörung der Bindungsmatrix von SiO_2 . Es wird jedoch die lokale Bindungsstruktur verändert. Dies kann benachbarte Si-Si-Bindungen beeinflussen und somit indirekt einen Betrag zum dielektrischen Durchbruch liefern. Die zweite Annahme ist bisher nicht belegt, kann jedoch die Differenz zwischen dem E-Modell und den Daten in Abbildung 5.23 erklären. Wenn ein HBD durch Wasserstoff beeinflusst wird, dann würde ein Experiment die etwas höhere Feldbeschleunigung messen, die auch für den SBD ermittelt wird (vgl. Abbildung 5.23). Im Gegensatz dazu wird für den HBD die höhere Aktivierungsenergie gemessen, die der Si-Si-Bindungsenergie entspricht. Für den SBD wird die niedrigere Aktivierungsenergie gemessen, welche zwischen den Bindungsenergien von H-Si und H-O liegt (vgl. Abbildung 5.21).

Die Modellvorstellung für den SBD stimmt mit den experimentellen Voraussetzungen überein. Die experimentellen Daten von Kimura wurden für ein 9,9 nm dickes Na_2O -Oxid ermittelt [5.36]. Damit ist der Einfluß von Wasserstoff im SiO_2 auf den HBD möglich. Wie bereits erwähnt, wurden die eigenen experimentellen Daten für ein nitridiertes Oxid gemessen. Vor der Nitridierung erfolgte eine Trockenoxidation und eine Na_2O -Oxidation des Siliziumsubstrates.

Analogie zum thermochemischen E-Modell

Die experimentellen Daten für den SBD zeigen eine Analogie zum thermochemischen E-Modell, das einen dielektrischen Durchbruch (HBD) aufgrund des Aufbrechens von Si-Si-Bindungen postuliert. Das Modell basiert auf der Verkleinerung der Aktivierungsenergie für das Aufbrechen einer Si-Si-Bindung durch das elektrische Feld. Dies erfolgt durch die Tatsache, daß die Dipolmomente der beiden $(\text{O}_3\equiv\text{Si})$ -Anteile in einem $(\text{O}_3\equiv\text{Si}-\text{Si}\equiv\text{O}_3)$ -Komplex im lokalen elektrischen Feld entgegengesetzt orientiert sind. Die Si-Si-Bindung wird daher mit zunehmendem elektrischen Feld stärker „belastet“. Die potentielle Energie, welche die Aktivierungsenergie für das Aufbrechen einer Si-Si-Bindung reduziert, ist durch

das Produkt aus $(\text{O}_3\equiv\text{Si})$ -Dipolmoment und lokalem elektrischen Feld bestimmt. Eine entsprechende Situation liegt bei einem $(\text{O}_3\equiv\text{Si}-\text{H})$ -Komplex bzw. einem $(\text{O}_3\equiv\text{Si}-\text{O}-\text{H})$ -Komplex nicht vor. Der Gradient des elektrischen Feldes bewirkt jedoch ein Drehmoment auf einen Dipol, der nicht entlang einer Äquipotentiallinie liegt. Die potentielle Energie des Dipols im elektrischen Feld ist das Vektorprodukt aus Dipolmoment und elektrischem Feld. Die Abnahme der Aktivierungsenergie für die H-Si-Bindung und die H-O-Bindung ist daher ebenfalls durch das Produkt aus dem jeweiligen Dipolmoment und dem lokalen elektrischen Feld bestimmt.

Zusammenhang zwischen SBD und SILC

Im Kapitel 5.1.1 wurde gezeigt, daß unterhalb eines Gatespannungsbetrages von ca. 5 V kein meßbarer SILC generiert wird. Diese Grenze kann mit einer Energie von 2 eV für das „Hydrogen Release Model“ in Verbindung gebracht werden und in der Tat generieren nur Elektronen mit einer potentiellen Energie von mehr als ca. 2 eV „Stress-Induced Leakage Current“ (SILC) [5.9]. Aus dem Verlauf der Aktivierungsenergie für den SBD, der in der Abbildung 5.22 gezeigt ist, könnte geschlossen werden, daß sich bei ca. 5 V der Mechanismus ändert, welcher zu SBD führt, und oberhalb von 5 V mit dem Mechanismus für SILC übereinstimmt. Die in diesem Kapitel dargelegte Modellvorstellung für den SBD stellt die Abnahme der Aktivierungsenergie jedoch in den Zusammenhang mit dem Einfluß des elektrischen Feldes auf H-Si-Bindungen und H-O-Bindungen. Ein Zusammenhang zwischen SBD und SILC wäre dennoch gegeben, wenn der Wasserstoff und dieselben Haftstellen, die einen leitenden SBD-Pfad bilden, auch für den SILC erforderlich sind. Im Kapitel 4.4.2 wurde aus experimentellen Ergebnissen gefolgert, daß die unterschiedlichen Gateleckströme des SILC und nach SBD mit unterschiedlichen Ionisationsenergien der generierten Haftstellen in Verbindung gebracht werden können.

5.4.4 Auswirkungen auf die Zuverlässigkeitscharakterisierung

Zunächst sei erwähnt, daß der lokale Charakter des „Soft Breakdown“ (SBD) sowie der statistische Prozeß der Generation von Haftstellen im Oxid eine Verwendung der Poisson-Statistik rechtfertigen. Damit kann für den SBD dieselbe statistische Datenauswertung verwendet werden wie für den dielektrischen Durchbruch (HBD). Insbesondere gilt auch für den SBD der im Kapitel 2.3.4 diskutierte Zusammenhang zwischen der Zeit bis zum Durchbruch und der Gatefläche.

Die gemessenen Zeiten bis zum SBD sind natürlich kleiner als die Zeiten bis zum HBD. Die Lebensdauer eines Oxides wird aber entscheidend durch die Feldbeschleunigung bestimmt. Da die Feldbeschleunigung für den SBD ebenfalls kleiner sein kann als für den HBD, ergeben sich im Fall des SBD deutlich kleinere intrinsische Lebensdauern. Aufgrund des im Kapitel 5.2 vorgestellten und diskutierten Modells ist dies für den NMOSFET, nicht aber für den PMOSFET relevant. Da der SBD den NMOSFET nur degradiert, wenn er im Überlapp von Gate und Drain auftritt, ist nur ein Teil der aktiven Gatefläche betroffen. Als Beispiel sind in der Abbildung 5.24 Daten des 3,4 nm dicken Oxides extrapoliert. Zuerst erfolgte eine Projektion von der Streßspannung $IV_g| = 5 \text{ V}$ auf eine Betriebsspannung von 1,8 V nach dem E-Modell. Anschließend wurde die Umrechnung von 125°C auf 80°C sowie von 0,0004 cm²

auf $0,1 \text{ cm}^2$ für den HBD und $0,001 \text{ cm}^2$ für den SBD vorgenommen. Die gemessenen Daten, welche in der Abbildung 5.24 links eingetragen sind, weisen keine extrinsischen Fehler auf. Der Punkt mit dem kleinsten Fehleranteil gibt die nachgewiesene minimale Defektdichte an. Im schlimmsten Fall geht die Fehlerverteilung bei diesem Punkt in den extrinsischen Anteil über. Von der Steilheit der extrinsischen Fehlergeraden des SBD und des HBD hängt es ab, ob das angestrebte Ziel einer bestimmten Lebensdauer mit einem bestimmten maximalen Fehleranteil für den NMOSFET erfüllt ist.

In der Abbildung 5.24 sind noch nicht die besonderen Struktur- und Potentialverhältnisse an der Drainkante für den SBD berücksichtigt. Im Überlappbereich von Gate und Drain ist der Vogelschnabel des Gateoxides lokalisiert („gate birds beak“). Das dickere Gateoxid führt an dieser Stelle zu einer Verkleinerung des elektrischen Feldes. Angenommen die Oxiddicke sei gleich, dann kann jedoch im Off-Zustand das elektrische Feld im Überlappbereich von Drain und Gate größer sein als in der aktiven Fläche über dem Substrat. Dies ist der Fall, wenn das Drain-Gebiet an der Grenzfläche zum Gateoxid nicht vollständig invertiert ist und ergibt sich aus der unterschiedlichen Austrittsarbeitendifferenz von n-dotiertem Drain und p-dotierter Wanne. In der Abbildung 5.25 sind die Potentialverhältnisse für das p-Substrat und das n-dotierte „Lightly Doped Drain“-Gebiet (LDD) miteinander verglichen.

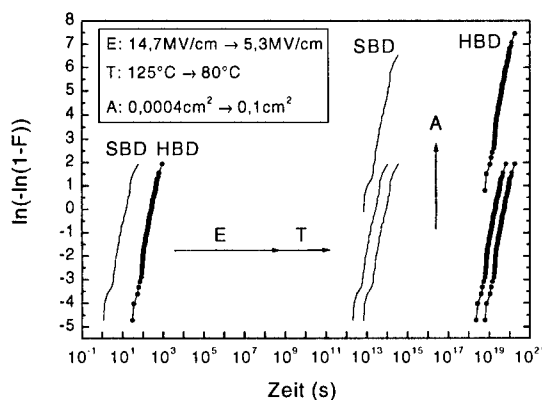


Abb. 5.24: Beispiel für eine Extrapolation der Lebensdauer auf Betriebsbedingungen. Es ist die Projektion des elektrischen Feldes E nach dem E -Modell, die Projektion der Temperatur T und die Projektion der Fläche A gezeigt. Der kleinere Feldbeschleunigungsfaktor für SBD führt zu einer deutlich kleineren intrinsischen Lebensdauer. Die Oxiddicke ist 3,4 nm.

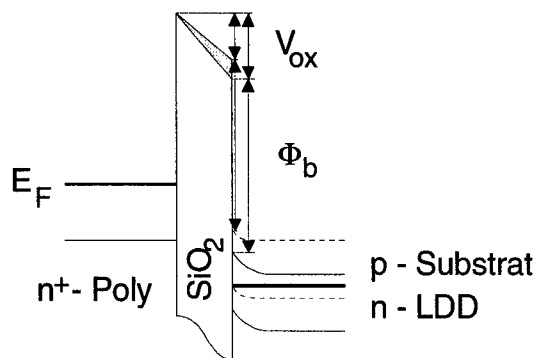


Abb. 5.25: Energiediagramm für den Off-Zustand eines NMOSFETs. An der Drainkante tritt bei gleicher Oxiddicke ein höheres elektrisches Feld auf als in der Gatefläche über dem p-Substrat. Als Anode sind das p-Substrat und ein n-dotiertes LDD-Gebiet („Lightly Doped Drain“) skizziert. Die Barrierenhöhe zwischen den Leitungsbandkanten von SiO₂ und Silizium ist in beiden Fällen gleich. Das Fermi-niveau E_F ist jeweils durch eine dickere Linie gekennzeichnet.

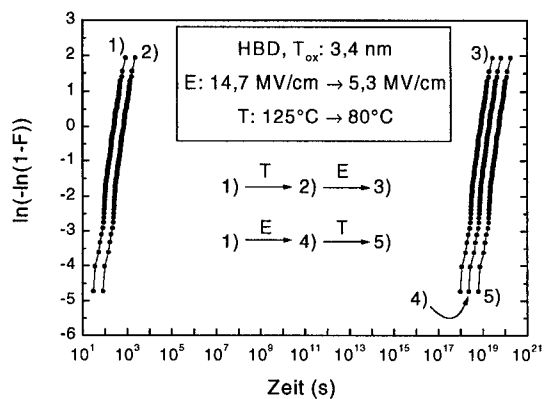


Abb. 5.26: Unterschiedliche Reihenfolge der Temperaturprojektion und der Feldprojektion für die Extrapolation der Lebensdauer. Die Abhängigkeit des Feldbeschleunigungsfaktors von der Temperatur ergibt eine Differenz von ca. einer Größenordnung in der Zeit. Die Meßdaten für die Zeit bis zum HBD bei 125°C und 14,7 MV/cm sind mit 1) gekennzeichnet.

Liegt eine temperaturabhängige Feldbeschleunigung vor, dann kann die Extrapolation für die Temperatur und das elektrische Feld nicht getrennt voneinander erfolgen. Die getrennte Berechnung wird abhängig von der Reihenfolge. In der Abbildung 5.26 ist dies am Beispiel der Meßdaten für das 3,4 nm dicke Oxid und den HBD gezeigt. Die Differenz für die unterschiedlichen Reihenfolgen beträgt in diesem Beispiel ca. eine Größenordnung in der Zeit.

5.5 Zusammenfassung und Ausblick auf die zukünftigen Prozeßgenerationen

Zu Beginn des Kapitels wurden die grundlegenden Charakteristika von „Stress-Induced Leakage Current“, „Soft Breakdown“ und dielektrischem Durchbruch („Hard Breakdown“) zusammengefaßt. Dabei wurde eine Meßroutine zur Bestimmung der Zeit bis zum „Soft Breakdown“ vorgestellt, die sehr einfach in standardisierte Streßmessungen bei konstanter Spannung implementiert werden kann. Zudem wurde gezeigt, daß der dielektrische Durchbruch nicht auf einen Ort in der Gatefläche einer MOS-Kapazität beschränkt ist. Nach einer Erhöhung der elektrischen Feldstärke treten mehrere dielektrische Durchbrüche, verteilt über die gesamte Gatefläche, auf. Es wurde ausführlich dargelegt, warum für Oxide dünner als ca. 5 nm ein elektrischer Streß bei konstanter Spannung gegenüber einem elektrischen Streß bei konstanter Stromdichte zu bevorzugen ist. Bei Oxiden dicker als ca. 8 nm ist aufgrund der Generation von Oxidladungen ein Streß bei konstanter Stromdichte anzuwenden, um das elektrische Feld im Oxid konstant zu halten.

Es wurde erstmals gezeigt, daß der „Soft Breakdown“ zu einer unmittelbaren Degradation von NMOS-Feldeffekttransistoren führen kann. Diese tritt auf, wenn der Ort des „Soft Breakdown“ im Überlappbereich von Gate und Drain lokalisiert ist. Die Degradation äußert sich in einem stark erhöhten Off-Strom des Transistors, der durch einen erhöhten „Gate-Induced Drain Leakage“-Strom verursacht wird. Ein Modell zur Erklärung dieses Effektes wurde dargelegt. Dieses basiert auf der Anreicherung von negativen Ladungen im Bereich des „Soft Breakdown“. Das Modell begründet auch, warum eine vergleichbare Degradation im Fall eines PMOS-Feldeffekttransistors nicht gemessen wird. Ausgehend von diesem Modell wurde für den „Soft Breakdown“ das Modell eines selbstlimitierenden Strompfades für den Transport von Elektronen entwickelt. Dieses Modell erklärt die Fluktuationen im Gateleckstrom nach einem „Soft Breakdown“ und konnte durch experimentelle Ergebnisse untermauert werden.

Die Beschleunigung der elektrischen Degradation dünner Oxide wurde in Abhängigkeit vom elektrischen Feld und von der Temperatur analysiert. Dabei erfolgte eine getrennte Charakterisierung für den „Soft Breakdown“ und für den dielektrischen Durchbruch. Es wurde gezeigt, daß sich das Beschleunigungsverhalten des „Soft Breakdown“ deutlich von dem des dielektrischen Durchbruchs unterscheiden kann. Damit steht fest, daß beide Durchbruchereignisse stets getrennt voneinander charakterisiert werden müssen. Die Auswirkungen der Degradation des NMOS-Feldeffekttransistors durch „Soft Breakdown“ auf

die Charakterisierung der Zuverlässigkeit von dünnen Gateoxiden wurde ausführlich diskutiert.

Die experimentellen Daten für den „Soft Breakdown“ (SBD) zeigen eine Analogie zum thermochemischen E-Modell. Aufgrund dieser Analogie wurde eine Modellvorstellung für die Ausbildung eines SBD-Pfades entwickelt. Diese legt das Aufbrechen von H-Si-Bindungen und H-O-Bindungen unter dem Einfluß des elektrischen Feldes zugrunde. Die Fragmente bilden die Haftstellen für den Ladungsträgertransport. Eine Anreicherung von negativen Ladungen im Bereich des „Soft Breakdown“ bzw. im SBD-Pfad kann mit dieser Modellvorstellung erklärt werden, wenn die Emission von Elektronen durch Si \cdot - und O \cdot - eine geringere Reaktionsrate aufweist, als das Besetzen von H $^+$, Si \cdot und O \cdot mit einem Elektron.

Die Qualifikation der Gateoxide der nächsten Prozeßgenerationen ist durch einen elektrischen Streß im Bereich des direkten Tunnelns geprägt. Dabei müssen alle Degradationsphänomene berücksichtigt werden. Bisher ist jedoch noch nicht endgültig geklärt, ob die Lebensdauer von SiO $_2$ -Schichten mit Dicken kleiner als ca. 2,0 nm noch durch einen vergleichbaren dielektrischen Durchbruch limitiert ist, wie dies bei dicken Oxiden der Fall ist. In der Literatur werden Versuche unternommen, ein Skalierungslimit für die SiO $_2$ -Schichtdicke zu erhalten, das im wesentlichen auf Aussagen über die Zuverlässigkeit dieser Schichten beruht. Unabhängig von einem solchen Skalierungslimit ist in den nächsten Jahren die minimal mögliche Dicke erreicht, die aus physikalischer Sicht eine SiO $_2$ -Schicht darstellt. Bei einer weiteren Skalierung der Prozeßgenerationen wird die Einführung von neuen Gatedielektrika notwendig. Der direkte Tunnelstrom bleibt bis zu dieser Dickengrenze bei der entsprechenden Skalierung der Gatelänge kleiner als der Drainstrom der MOS-Feldeffekttransistoren. Dies ist der Fall, da die Gatefläche proportional mit der Gatelänge abnimmt, der Drainstrom jedoch proportional mit der Gatelänge zunimmt [5.46].

6 Titandioxid als alternatives Gatedielektrikum

Das klassische Gateoxid in MOS-Feldeffekttransistoren wird in der Halbleiterprozeßtechnik als erste Schicht die physikalisch realisierbaren Grenzen erreichen. Ab Technologien mit einer minimalen Strukturlänge von weniger als 100 nm wird die Einführung von Gatedielektrika mit höheren Dielektrizitätskonstanten erforderlich. Siliziumnitrid wird als ein Kandidat zur Verwendung als Gatedielektrikum in Erwägung gezogen [6.1, 6.2, 6.3]. Die relative Dielektrizitätskonstante von Siliziumnitrid beträgt 7,9 [6.4]. Damit läßt sich gegenüber Siliziumdioxid, das eine relative Dielektrizitätskonstante von 3,9 besitzt, ein Faktor zwei in der Schichtdicke erzielen. Höhere Dielektrizitätskonstanten erfordern die Verwendung von Metalloxidschichten. Aktuelle Arbeiten beschäftigen sich u.a. mit dem aus der DRAM-Technologie bekannten Material Tantalpentoxid [6.5, 6.6, 6.7, 6.8], mit Aluminiumoxid [6.9, 6.10] und mit Titandioxid [6.11, 6.12]. In diesem Kapitel wird einer der ersten Versuche vorgestellt, Titandioxid in das Gate eines kompletten CMOS-Prozesses zu integrieren. Während der Charakterisierungsphase wurde ein ähnlicher Versuch beim Symposium on VLSI Technology 1999 vorgetragen [6.13].

6.1 Eigenschaften und Abscheidung von Titandioxid

6.1.1 Materialeigenschaften

Titandioxid kristallisiert in drei polymorphen Formen, die mit Anatase, Brookite und Rutil bezeichnet werden [6.14]. Anatase und Brookite transformieren bei 700°C bzw. 900°C in die Rutil-Phase [6.14]. Die Kristallisationstemperatur von Anatase liegt unter 500°C [6.15]. In der Tabelle 6.1 sind wichtige physikalische Eigenschaften der Rutil-Phase zusammengefaßt. Es gibt in der Literatur kontroverse Aussagen über direkte und indirekte Bandlücken von TiO_2 [6.14].

6.1.2 Abscheideverfahren

Für die Abscheidung von Titandioxid sind in der Literatur vier Verfahren bekannt. Die MOCVD-Abscheidung („Metal Organic Chemical Vapor Deposition“) ist mit verschiedenen Trägergasen wie z.B. „titanium tetrakis-isopropoxide“ [6.16] oder „tetra-isopropyl-titanate“ [6.17] durchführbar. Die Abscheidetemperatur beträgt ca. 300°C bis 350°C. Zur Verbesserung der Schichteigenschaften ist jedoch nach der Abscheidung eine Reoxidation bei bis zu 800°C erforderlich. Eine andere Möglichkeit bietet die thermische Oxidation von Titanitrid. Die Abscheidung von Titanitrid kann entweder mittels eines MOCVD-Verfahrens, mittels

Sputtern oder mittels eines PECVD-Verfahrens („Plasma Enhanced CVD“) erfolgen. Titandioxid kann auch durch reaktives Sputtern von Titan abgeschieden werden. Das Titandioxid kristallisiert dabei als Rutile [6.15]. Das vierte Verfahren wird als „Jet Vapor Deposition“ (JVD) bezeichnet. Eine spezielle Anlage ermöglicht in diesem Verfahren die Schichtabscheidung bei niedrigen Temperaturen mit Hilfe von gerichteten Strahlen aus Atomen, hier also Titanatomen und Sauerstoffatomen. Nach der Abscheidung der Schicht ist eine Reoxidation bei ca. 600°C bis 750°C erforderlich [6.12].

Konstanten des tetragonalen Gitters bei 30°C	$a = 4,59 \text{ \AA}, c = 2,96 \text{ \AA}$
Dielektrizitätskonstante: parallel zu a	$\epsilon = 86 \cdot \epsilon_0$
parallel zu c	$\epsilon = 170 \cdot \epsilon_0$
Betrag der Bandlücke	zwischen 3,0 eV und 3,1 eV
Linearer thermischer Ausdehnungskoeffizient	$\alpha_{ } = 8,816 \cdot 10^{-6} \text{ K}^{-1}$ $\alpha_{\perp} = 7,249 \cdot 10^{-6} \text{ K}^{-1}$

Tab. 6.1: Wichtige physikalische Eigenschaften der Rutile-Phase von Titandioxid [6.14].

Die beschriebenen Verfahren weisen den Nachteil auf, daß ein CVD-Prozeß involviert ist oder spezielle Anlagen erforderlich sind. CVD-Schichten enthalten immer einen Anteil an Fremdatomen und besitzen schlechtere Eigenschaften in bezug auf Defektdichte und Grenzflächen. Die MOCVD-Prozesse erfordern bestimmte Anlagen bzw. Gasversorgungen und das JVD-Verfahren basiert gänzlich auf einer speziellen Prozeßanlage. Das reaktive Sputtern von Titan benötigt ein separates Sputtertarget, um Standardprozesse nicht nachhaltig zu beeinflussen. Es wurde daher ein neues Verfahren eingeführt, das ausschließlich anhand von Standardprozessen realisierbar ist [6.18]. Das Verfahren basiert auf der thermischen Oxidation von gesputtertem Titan. Sowohl das Sputtern von Titan als auch die thermische Oxidation in einer RTP-Anlage („Rapid Thermal Process“) wurden mit Hilfe von Standardprozessen durchgeführt. Weitere Vorteile des neuen Verfahrens sind die Reinheit der Titandioxidschicht und der hohe Oxidationsgrad aufgrund der stark oxidierenden Eigenschaft von Titan. Im Gegensatz zu den oben beschriebenen Verfahren ist eine Barrierschicht zwischen Siliziumsubstrat und gesputterter Titanschicht erforderlich, um die Bildung von Titansilizid zu vermeiden. Bei den MOCVD-Verfahren bildet sich jedoch in der sauerstoffreichen Atmosphäre eine ungewollte SiO_2 -Schicht zwischen dem Siliziumsubstrat und der Titandioxidschicht aus.

Generell ist bei nachfolgenden Prozessen, z.B. bei der Abscheidung von Polysilizium, eine Kontamination durch Titan zu vermeiden. Titandioxid ist in einem weiten Temperaturbereich zwischen 400°C und 1700°C stabil [6.14]. Daher ist keine Kontamination von Prozeßanlagen durch Titanatome zu erwarten.

6.2 Prozeßkonzepte mit Titandioxid als Gatedielektrikum

Ziel der Verwendung alternativer Dielektrika ist es, ein Dielektrikum zu implementieren, dessen Kapazität diejenige einer 1,5 nm dicken SiO₂-Schicht überschreitet. Die Schichtdicke des neuen Dielektrikums ist so groß zu wählen, daß direktes Tunneln von Elektronen vermieden wird. Zudem sollte der natürliche Leckstrom des Dielektrikums deutlich kleiner sein, als der direkte Tunnelstrom durch eine SiO₂-Schicht mit einer Dicke von 1,5 nm oder weniger. Aufgrund des kapazitiven Vergleichs ist die Einführung einer effektiven Oxiddicke T_{eff} sinnvoll, die folgendermaßen definiert wird:

$$\frac{1}{C_{\text{DIE}}} = \sum_i \frac{1}{C_i} = \frac{T_{\text{DIE}}}{\epsilon_{\text{DIE}}} \quad (6.1)$$

$$T_{\text{eff}} = \frac{\epsilon_{\text{ox}}}{\epsilon_{\text{DIE}}} \cdot T_{\text{DIE}} = \frac{\epsilon_{\text{ox}}}{C_{\text{DIE}}}$$

Dabei sind C_{DIE} die gesamte Flächenkapazität des Dielektrikums, C_i die Flächenkapazitäten der einzelnen Schichtanteile, T_{DIE} die gesamte Schichtdicke des Dielektrikums und ε_{DIE} die effektiv wirksame Dielektrizitätskonstante des Dielektrikums. ε_{ox} ist die Dielektrizitätskonstante von SiO₂ und beträgt 3,9·ε₀. Die Schicht mit der kleinsten Kapazität bestimmt die gesamte Kapazität des Dielektrikums.

6.2.1 Basisprozeß

Der Basisprozeß sieht eine Barrierschicht aus Siliziumdioxid vor, auf der Titan gesputtert wird. Nach der thermischen Oxidation in einer RTP-Anlage erfolgt die Abscheidung von Polysilizium. Dabei wird das Polysilizium an der Grenzfläche zum Titandioxid oxidiert. Diese SiO₂-Schicht wird im folgenden als „Topoxid“ bezeichnet. Die Oxidation des Polysiliziums erfolgt zunächst aufgrund einer Übersättigung von Titandioxid mit Sauerstoff. Im Kapitel 6.4 wird gezeigt, daß eine weitere Oxidation der Polysiliziumschicht durch die Temperaturschritte im nachfolgenden Prozeßverlauf verursacht wird. In der Abbildung 6.1 sind die Prozeßfolge und die Schichtfolge des Basisprozesses skizziert. Der Basisprozeß hat zunächst den Vorteil, daß die Änderungen am ursprünglichen CMOS-Prozeß minimiert sind. Die gesamte SiO₂-Schichtdicke aus Barrierschicht und Topoxid bestimmt jedoch die Gesamtkapazität des Dielektrikums.

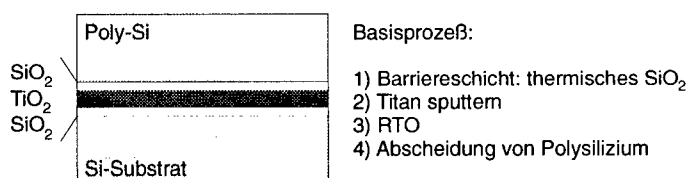


Abb. 6.1: Prozeßfolge und Schichtfolge des Basisprozesses

6.2.2 Prozeß mit Nitridschichten

Eine Erhöhung der Gesamtkapazität des Dielektrikums wird erreicht, wenn anstelle von SiO_2 jeweils Siliziumnitridschichten verwendet werden. In diesem Prozeßkonzept wird die Barrierschicht aus thermisch nitridiertem Silizium gebildet. Anstelle der thermischen Nitridierung von Silizium [6.19] kann aber auch ein sehr dünnes CVD-Nitrid verwendet werden. Vor der Abscheidung des Polysiliziums wird eine Schicht aufgebracht, welche die Oxidation des Polysiliziums verhindert. Eine solche Schicht wird im folgenden als „Oxidationssperrschicht“ bezeichnet. In diesem Prozeßkonzept ist als Oxidationssperrschicht ein CVD-Siliziumnitrid integriert. Die Si_3N_4 -Schicht soll auch die Oxidation des Polysiliziumgates während des weiteren Prozeßverlaufes verhindern. In der Abbildung 6.2 sind die Prozeßfolge und die Schichtfolge dieses Prozeßkonzeptes skizziert.

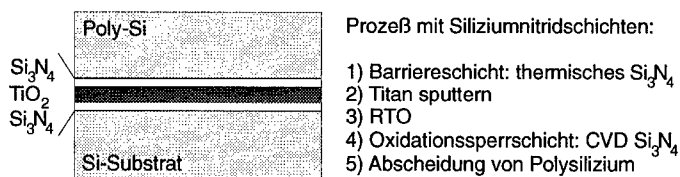


Abb. 6.2: Prozeßfolge und Schichtfolge des Prozeßkonzeptes mit Siliziumnitridschichten

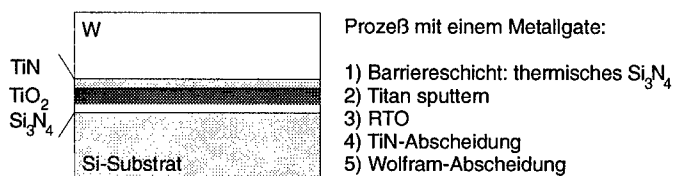


Abb. 6.3: Prozeßfolge und Schichtfolge des Prozeßkonzeptes mit einem Metallgate

6.2.3 Prozeß mit einem Metallgate

Eine weitere Verbesserung hinsichtlich der effektiven Oxiddicke ist möglich, wenn anstelle der oberen Siliziumnitridschicht eine leitende Oxidationssperrschicht verwendet wird. Die Kapazität wird dann nur noch durch die Barrierschicht aus Siliziumnitrid bestimmt. Eine Dicke der Siliziumnitridschicht von 2 nm würde also eine effektive Oxiddicke von ca. 1 nm ergeben. Als leitende Oxidationssperrschicht sind z.B. Titannitrid oder Wolframnitrid

denkbar. Damit ist dieser Prozeß für die Implementierung eines Metallgates geeignet. Prozeßkonzepte für Hochfrequenz-CMOS beinhalten die Integration eines Metallgates anstelle eines Polysiliziumgates, um den spezifischen Widerstand zu senken und damit die Schaltgeschwindigkeit zu erhöhen. In der Abbildung 6.3 sind die Prozeßfolge und die Schichtfolge des Prozeßkonzeptes mit einem Metallgate aus Wolfram und einer leitenden Oxidationssperrschicht aus Titannitrid skizziert.

6.3 Prozeßtests und Schichtanalysen

Für die ersten Prozeßtests und Schichtanalysen wurden unstrukturierte Wafer hergestellt. Der Prozeßablauf entspricht dem Basisprozeß. In der Abbildung 6.4 ist die Schichtfolge anhand einer Aufnahme an einem Transmissionselektronenmikroskop (TEM) gezeigt. Zwischen dem Siliziumsubstrat, das mit „Si“ bezeichnet ist, und der TiO_2 -Schicht ist die Barrierschicht aus Siliziumdioxid zu erkennen. Zwischen der TiO_2 -Schicht und der Polysiliziumschicht befindet sich die dünne Topoxidschicht.

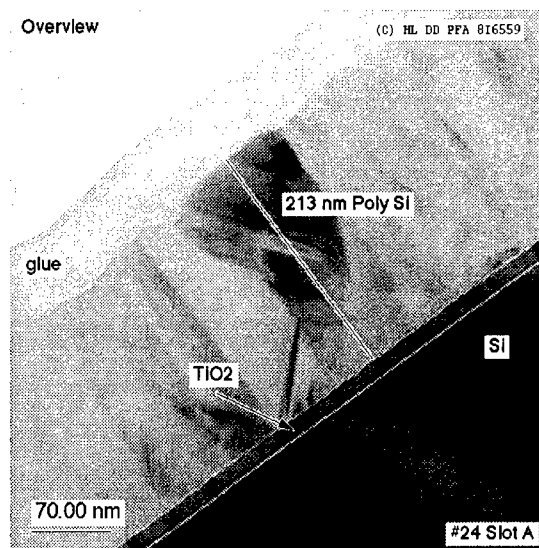


Abb. 6.4: TEM-Aufnahme eines Prozeßtests. Es ist die Schichtfolge des Basisprozesses zu erkennen. Oberhalb und unterhalb der TiO_2 -Schicht befinden sich jeweils Schichten aus Siliziumdioxid.

Die Abbildung 6.5 zeigt eine TEM-Aufnahme eines Prozeßtests, für den eine Barrierschicht aus ca. 3 nm thermisch gewachsenen SiO_2 verwendet wurde. Die Schichtdicke des gesputterten Titans war ca. 10 nm. Die thermische Oxidation von Titan ergibt eine ca. 14,3 nm dicke TiO_2 -Schicht. Damit entspricht die Vergrößerung der physikalischen Schichtdicke erwartungsgemäß etwa einem Faktor 1,5. Die Barrierschicht beträgt hier nach dem Aufbau der Gateschichten nur noch ca. 1,5 nm. Zwischen Titandioxid und Polysilizium ist erneut die Ausbildung der Topoxidschicht sehr gut zu erkennen, die hier ca. 0,8 nm dick ist.

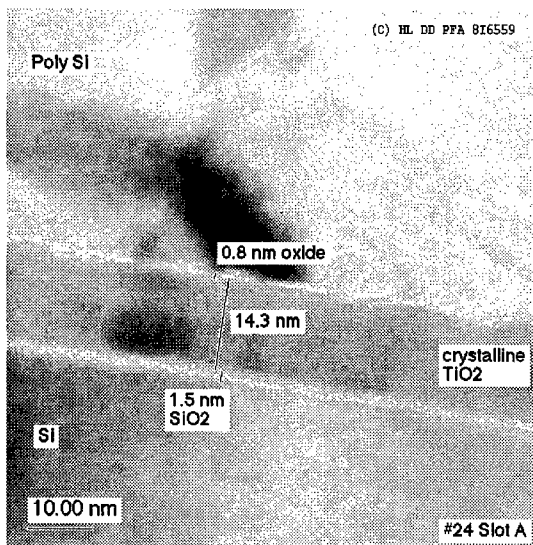


Abb. 6.5: TEM-Aufnahme eines Prozeßtests mit 10 nm gesputtertem Titan. Die SiO_2 -Barrierschicht war ursprünglich ca. 3 nm dick.

Augeranalysen weisen eindeutig die Ausbildung einer TiO_2 -Schicht nach. In der Abbildung 6.6 sind die Signalverläufe von Titan, Silizium und Sauerstoff in Abhängigkeit von der Sputterzeit gezeigt, die einer bestimmten Probentiefe entspricht. Anhand der beiden Signale des oxidisch gebundenen Siliziums bei 84 eV und des metallisch gebundenen Siliziums bei 95 eV kann die SiO_2 -Barrierschicht zwischen Substrat und Titandioxid gezeigt werden. Das Sauerstoffsignal für die Si-O-Bindung verhält sich entsprechend. Bei der Probenpräparation wurde die Polysiliziumschicht entfernt, so daß zu Beginn die Topoxidschicht und anschließend die Titandioxidschicht analysiert wurden. Das Titansignal im Bereich der Barrierschicht und des Siliziumsubstrates wird durch den Sputterprozeß verursacht. Eine Probe, für welche die Titandioxidschicht vor der Augeranalyse in der Probenpräparation entfernt wurde, zeigte in der Barrierschicht und im Siliziumsubstrat kein Titansignal.

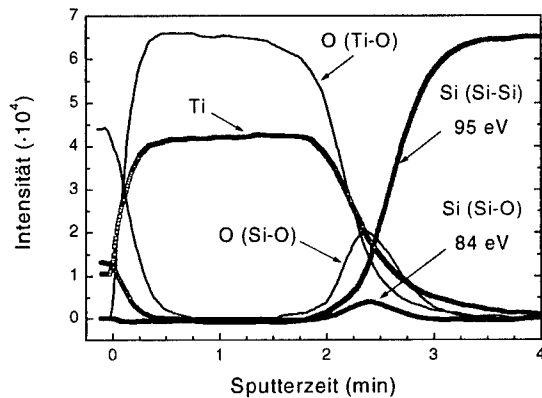


Abb. 6.6: Augertiefenprofil eines Prozeßtests. Es sind die Topoxidschicht, die Titandioxidschicht, die SiO_2 -Barrierschicht und das Siliziumsubstrat zu erkennen.

6.4 Implementierung in einen $0,25\mu\text{m}$ -CMOS-Prozeß

Um die notwendigen Prozeßänderungen zu minimieren, wurde der Basisprozeß in den vollständigen $0,25\mu\text{m}$ -CMOS-Prozeß integriert. Als Barrierschicht wurde eine ca. $1,5\text{ nm}$ dicke SiO_2 -Schicht verwendet. Die Schichtdicke des gesputterten Titan war ca. 5 nm . In den folgenden Kapiteln wird das Gatedielektrikum, das nun auch als $\text{SiO}_2/\text{TiO}_2$ -Dielektrikum bezeichnet wird, elektrisch und strukturell charakterisiert.

6.4.1 Physikalische Analyse der Schichtdicken und der Gatekante

Die Charakterisierung mittels Transmissionselektronenmikroskopie (TEM) zeigt, daß die Schichtdicken der Topoxidschicht und der SiO_2 -Barrierschicht während der weiteren Prozessierung nach der Gatestrukturierung zunehmen. In der Tabelle 6.2 sind die Schichtdicken aufgelistet, die anhand von TEM-Aufnahmen bestimmt wurden. Die Aufnahmen wurden in der Nähe der Gatekante und in der Mitte einer planaren Kapazität erstellt, die eine Länge und eine Breite von $0,1\text{ cm}$ bzw. $0,012\text{ cm}$ besitzt. Die Topoxidschicht und die Barrierschicht sind an der Gatekante deutlich dicker als in der Mitte der Kapazität. Die Schichtdicken in der Mitte der Kapazität sind wiederum dicker als in den Prozeßtests (vgl. Abbildung 6.5). Nach der Gatestrukturierung erfolgt demnach eine zusätzliche Oxidation, die nicht nur an der Gatekante stattfindet. Bei Reoxidationen im weiteren

Prozeßverlauf werden auch in einer Distanz von mindestens 0,012 cm ausreichend Sauerstoffatome für die weitere Oxidation zur Verfügung gestellt.

Schicht	Mittelwert in der Kapazitätsmitte	Mittelwert nahe der Gatekante
Topoxid (SiO_2)	1,5 nm	4,5 nm
TiO_2	9 nm	9 nm
Barrierschicht (SiO_2)	2,5 nm	4,5 nm

Tab. 6.2: Schichtdicken nach der vollständigen Prozessierung bis zur ersten Metallebene. Die Messungen erfolgten an TEM-Aufnahmen einer planaren Kapazität.

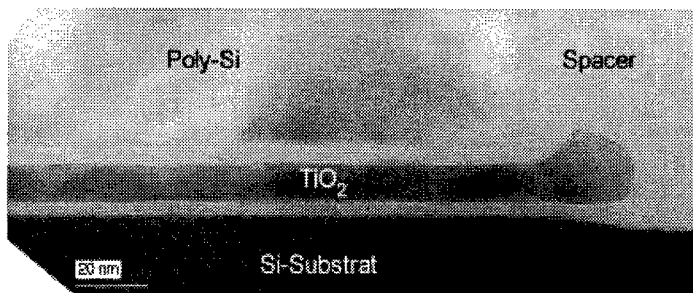


Abb. 6.7: TEM-Aufnahme an der Gatekante eines Kurzkanaltransistors. Das Titandioxid weitet sich im nachfolgenden Prozeßverlauf in den Bereich des Spacers aus.

Die Ätzung der Titandioxidschicht erfolgte unmittelbar nach der Strukturierung des Polysiliziumgates, um die nachfolgenden Implantationen des Standardprozesses verwenden zu können. Eine hochselektive Metalloxidätzung erlaubt die Strukturierung der Titandioxidschicht [6.20]. In der Abbildung 6.7 ist die Gatekante eines Kurzkanaltransistors mit $W/L = 10/0,25 \mu\text{m}$ gezeigt. Die Prozessierung nach der Gatestrukturierung führt zu einer Ausweitung des Titandioxides in den Bereich des Spacers. Über die Gatekante hinaus ergibt sich dadurch eine leichte Vergrößerung der Titandioxidschichtdicke, vergleichbar zum Vogelschnabel an der Gatekante eines SiO_2 -Gateoxides.

Die Abbildung 6.7 belegt zusätzlich, daß nach der vollständigen Prozessierung das Topoxid und die SiO_2 -Barrierschicht wesentlich dicker sind als in den Prozeßtests. Es ist damit die effektive Oxiddicke größer als erwartet.

6.4.2 Eingangskennlinien von Feldeffekttransistoren

In der Abbildung 6.8 sind die Eingangskennlinien von NMOS-Kurzkanaltransistoren mit $W/L = 10/0,20 \mu\text{m}$, dargestellt. Die NMOSFETs weisen einen deutlichen Unterschied zwischen Waferrand und zentralem Waferbereich auf. Die sehr hohe Einsatzspannung ist darauf zurückzuführen, daß die Kanalimplantation für eine deutlich kleinere effektive Oxiddicke justiert wurde. Die NMOSFETs im zentralen Bereich besitzen eine noch höhere Einsatzspannung und eine kleinere Steigung der Unterschwellgeraden. Beides deutet auf ein Dielektrikum mit noch größerer effektiven Oxiddicke T_{eff} hin.

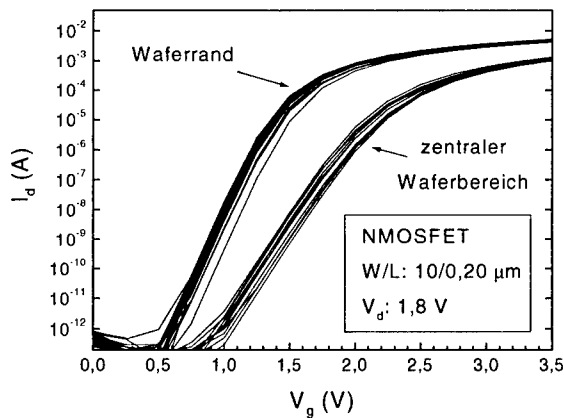


Abb. 6.8: Eingangskennlinien von NMOS-Kurzkanaltransistoren mit einem Dielektrikum aus SiO_2 -Schichten und einer TiO_2 -Schicht. Es existiert eine Verteilung von NMOSFETs mit höherer Einsatzspannung und kleinerer Steigung der Unterschwellgeraden.

In der Abbildung 6.9 sind die Eingangskennlinien von PMOS-Kurzkanaltransistoren gezeigt. Für die PMOSFETs existiert ebenfalls eine Verteilung von Transistoren, die eine größere Einsatzspannung und eine kleinere Steigung der Unterschwellgeraden aufweisen. Zudem besitzen alle Eingangskennlinien einen Kink-Effekt, der ein Hinweis auf einen parasitären Transistor mit einer kleineren Einsatzspannung ist. Da ein vergleichbarer Kink-Effekt bei den NMOSFETs nicht vorhanden ist, wird ausgeschlossen, daß dieser durch den Verlauf der Feldlinien oder durch die Kapazität des Gatedielektrikums an der Isolationskante verursacht wird. Im Fall von PMOS-Langkanaltransistoren, z.B. mit einer Gatelänge von $1,0 \mu\text{m}$, wird kein Kink-Effekt gemessen. In der Abbildung 6.7 wurde gezeigt, daß sich das Titandioxid über die Gatekante hinaus ausweitet. Dies ist auch an der Isolationskante der Fall. Es wird vermutet, daß die Ausweitung, welche während eines Reoxidationsprozesses stattfindet,

Punktdefekte im Substrat in Form von Zwischengitterplätzen verursacht. Für die thermische Oxidation von Silizium ist bekannt, daß die Ausbildung von Zwischengitterdefekten die Diffusion von Dotierstoffen erhöht oder erniedrigt. Dies wird als „oxidation-enhanced diffusion“ (OED) bzw. „oxidation-retarded diffusion“ (ORD) bezeichnet [6.21, 6.22]. Bor, Phosphor und Arsen zeigen eine erhöhte Diffusion. Im verwendeten CMOS-Prozeß wurden die Source- und Draingebiete des NMOSFETs mit Arsen und die des PMOSFETs mit Bor implantiert. Wenn durch die Ausweitung des Titandioxides an der Substratoberfläche Zwischengitterdefekte gebildet werden, dann diffundieren die Source- und Draingebiete des PMOSFETs lateral stärker aus. Damit erhöht sich der Anteil der Raumladungszonen der p-n-Übergänge gegenüber der Raumladungszone des Substrates an der Oberfläche. Die Kurzkanaleffekte, wie z.B. die Abnahme der Einsatzspannung mit der Kanallänge („ V_t roll-off“), werden also verstärkt. Das Kapitel 7.6 im Anhang A erläutert kurz die Ursache für die Kurzkanaleffekte. Die Dichte an Zwischengitterdefekten ist am größten, wo sich die Diffusionskante und die Isolationskante kreuzen. Der PMOSFET weist also an der Isolationskante einen stärkeren „ V_t roll-off“ auf und es ergibt sich ein parasitärer Transistor mit einer kleineren Einsatzspannung und einer kleineren Weite. Dies erklärt den Kink-Effekt im Fall von PMOSFETs. Im Fall von NMOSFETs ist die zusätzliche laterale Ausdiffusion von Source und Drain nicht so ausgeprägt, da die Aktivierungsenergie für die Diffusion von Arsen größer ist als für Bor. In der Literatur werden für die Bordiffusion unter intrinsischen Gleichgewichtsbedingungen Aktivierungsenergien zwischen 3,25 eV und 3,51 eV angegeben. Für Arsen existieren Angaben zwischen 4,08 eV und 4,23 eV [6.22].

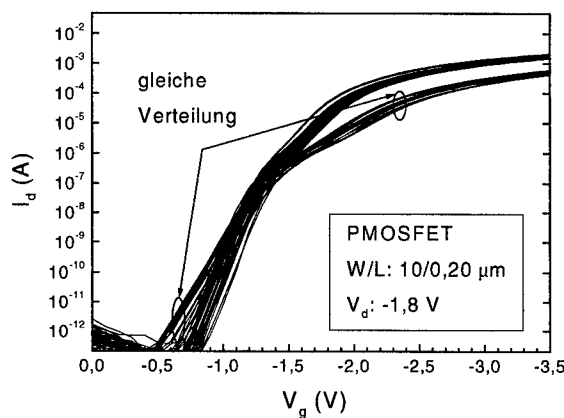


Abb. 6.9: Eingangskennlinien von PMOS-Kurzkanaltransistoren mit einem Dielektrikum aus SiO_2 -Schichten und einer TiO_2 -Schicht. Es existiert eine Verteilung von PMOSFETs mit höherer Einsatzspannung und kleinerer Steigung der Unterschwellgeraden. Zudem tritt in allen Kennlinien ein Kink-Effekt auf.

In der Abbildung 6.10 sind die Eingangskennlinien der NMOSFETs am Waferrand aus der Abbildung 6.8 mit denjenigen von NMOSFETs mit einem SiO_2 -Gateoxid verglichen. Die Einsatzspannung der NMOSFETs mit dem nominell 1,5 nm dicken Gateoxid ist etwas größer, da das Gateoxid eine elektrisch relevante Oxiddicke von ca. 2,1 nm besitzt. Die Eingangskennlinien der NMOSFETs mit einem $\text{SiO}_2/\text{TiO}_2$ -Gatedielektrikum weisen eine deutlich größere Streuung auf.

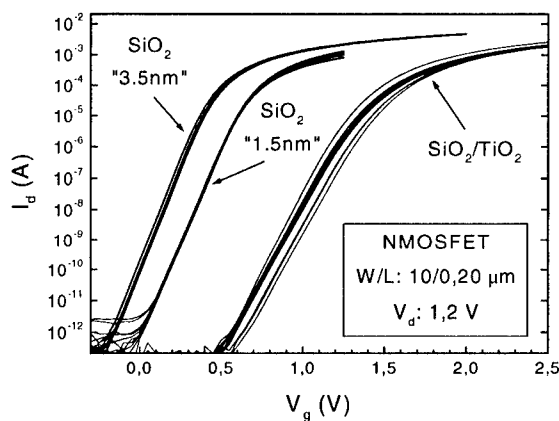


Abb. 6.10: Eingangskennlinien von NMOSFETs mit einem $\text{SiO}_2/\text{TiO}_2$ -Gatedielektrikum im Vergleich mit NMOSFETs mit einem SiO_2 -Gateoxid.

6.4.3 Leckströme des Dielektrikums

Die Charakterisierung des Leckstromes durch das $\text{SiO}_2/\text{TiO}_2$ -Gatedielektrikum erfolgt jeweils im direkten Vergleich mit einem 3,5 nm dicken SiO_2 -Gateoxid. Es werden immer NMOS-Kapazitäten bzw. NMOS-Transistoren bei positiven und negativen Gatespannungen miteinander verglichen. In der Abbildung 6.11 sind die $J_g(V_g)$ -Kurven gezeigt, die an einer großflächigen Kapazitätsstruktur gemessen wurden. Das $\text{SiO}_2/\text{TiO}_2$ -Dielektrikum besitzt in diesem Fall eine Leckstromdichte, die vergleichbar mit der Stromdichte nach einem dielektrischen Durchbruch ist. In der Abbildung 6.12 sind die Stromdichten gezeigt, die an einer Kammstruktur gemessen wurden, welche die gleiche gesamte Gatefläche besitzt, wie die großflächige Kapazität in der Abbildung 6.11. Die Kammstruktur besteht aus parallel kontaktierten Gestegen, die diffusionsbegrenzt sind. Die Stromdichte weist zwar im Fall der Kammstruktur eine weite Streuung auf, ist aber um Größenordnungen kleiner als diejenige im Fall der planaren Kapazität. Die Abbildung 6.13 zeigt die Gateleckstromdichten, welche an parallelen Kurzkanaltransistoren gemessen wurden, die insgesamt eine Gatefläche von $0,0009 \text{ cm}^2$ besitzen. Die gesamte Gatefläche ist also in etwa gleich groß wie die Gateflächen

der planaren Kapazität und der Kammstruktur. Die Leckstromdichte ist gegenüber der Kammstruktur nochmals um Größenordnungen kleiner und die Streuung ist ebenfalls geringer. Als letzter Vergleich sind in der Abbildung 6.14 die Gateleckstromdichten eines Kurzkanaltransistors mit einem $\text{SiO}_2/\text{TiO}_2$ -Dielektrikum und die Stromdichte durch das SiO_2 -Gateoxid gezeigt. Die Leckstromdichte des $\text{SiO}_2/\text{TiO}_2$ -Dielektrikums ist praktisch gleich groß wie bei den parallelen Transistoren.

Aus den Vergleichen für die verschiedenen Kapazitäten und Transistoren kann geschlossen werden, daß die Leckstromdichte durch das $\text{SiO}_2/\text{TiO}_2$ -Dielektrikum im Zusammenhang mit der mechanischen Streßrelaxation der Schicht steht. Die Kammstruktur ermöglicht eine wesentlich bessere Streßrelaxation als die planare Kapazität. Gegenüber der Kammstruktur sind bei der Struktur mit den parallelen Transistoren die Gatestege länger und nur ein Teil der gesamten Polysiliziumfläche trägt zur aktiven Gatefläche bei. Damit ist hier die Streßrelaxation relativ zur aktiven Gatefläche besser als bei der Kammstruktur. Der Einzeltransistor in der Abbildung 6.14 hat nur eine aktive Gatefläche von $2,5 \cdot 10^{-8} \text{ cm}^2$. Im Kapitel 6.4.4 wird gezeigt, daß die Titandioxidschicht große Körner ausbildet. Zwischen den Körnern können Inseln mit SiO_2 existieren. Es wird vermutet, daß diese SiO_2 -Inseln mit dem mechanischen Streß im weiteren Prozeßverlauf nach der Gatestrukturierung korreliert sind. Es ist davon auszugehen, daß das SiO_2 in diesen Inseln eine hohe Defektdichte aufweist, insbesondere an den vertikalen Grenzflächen zum TiO_2 . Diese hohe Defektdichte ist für einen sehr hohen Stromtransport verantwortlich, wie er in der Abbildung 6.11 gezeigt ist.

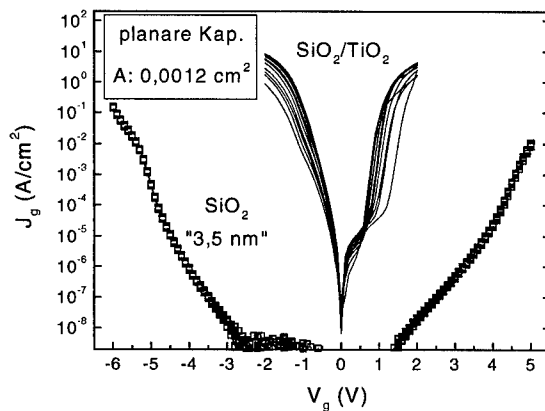


Abb. 6.11: Verteilung von Gateleckstromdichten, die jeweils an einer großflächigen planaren NMOS-Kapazität gemessen wurden.

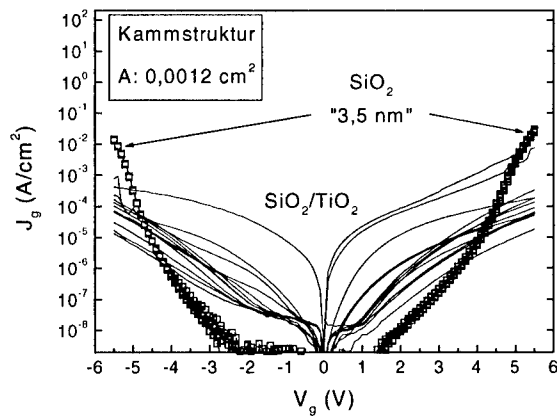


Abb. 6.12: Verteilung von Gateleckstromdichten, die jeweils an einer Kammstruktur gemessen wurden. Die Kammstruktur besitzt dieselbe gesamte Gatefläche, wie die planare Kapazität in der Abbildung 6.11.

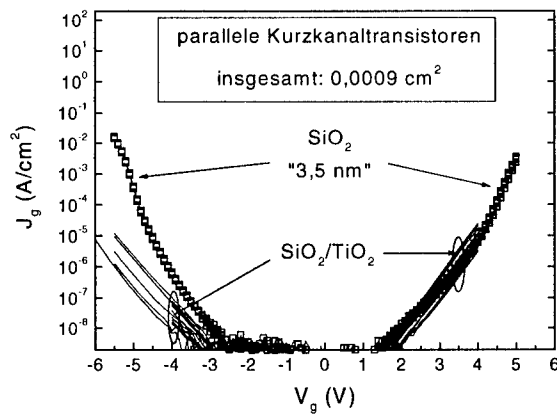


Abb. 6.13: Verteilung von Gateleckstromdichten, die jeweils an parallel geschalteten Transistoren gemessen wurden. Diese besitzen insgesamt eine Gatefläche von 0,0009 cm².

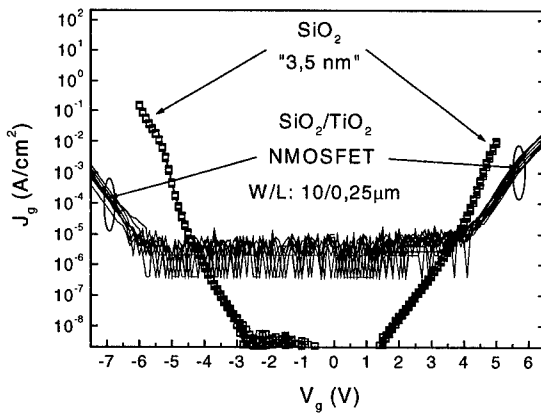


Abb. 6.14: Gateleckstromdichten im Fall eines Kurzkanaltransistors mit $\text{SiO}_2/\text{TiO}_2$ -Dielektrikum im Vergleich mit den Stromdichten für das SiO_2 -Gateoxid. Für die an den NMOSFETs gemessenen Stromdichten ist das Niveau des Grundrauschens hoch, da die gesamte Gatefläche sehr klein ist.

6.4.4 Strukturelle Schichteigenschaften des Dielektrikums

Die TEM-Aufnahme in der Abbildung 6.15 zeigt die Ausbildung von großen kristallinen TiO_2 -Körnern. Die Bereiche zwischen den einzelnen Titandioxidkörnern werden mit Siliziumdioxid gefüllt. Dies ist in der Abbildung 6.15 bei den Korngrenzen in der Nähe der Topoxidschicht und der Barrierschicht zu erkennen. Im Siliziumdioxid entlang diesen Korngrenzen ist eine sehr hohe Defektdichte zu erwarten.

In der Abbildung 6.16 ist die TEM-Aufnahme an einer sehr dünn präparierten Probe gezeigt. Die Aufnahme wurde in der Mitte einer großflächigen, planaren Kapazität gemacht. Es ist deutlich zu erkennen, daß hier Bereiche zwischen den TiO_2 -Körnern existieren, in denen sich ausschließlich Siliziumdioxid befindet. Dies erklärt die hohen Leckströme im Fall von planaren Kapazitäten. Im Gegensatz dazu zeigt die Abbildung 6.15 das Dielektrikum im Fall eines Kurzkanaltransistors. Die Lücken sind nicht so groß und die Leckstromdichte entsprechend klein, wie im Kapitel 6.4.3 gezeigt wurde. Im Kapitel 6.4.3 wurde der Zusammenhang des Leckstromes mit der mechanischen Streßrelaxation diskutiert. Die Abbildung 6.16 zeigt, daß der hierfür kritische Prozeßschritt vor oder während der Prozeßphase erfolgt, die zur weiteren Oxidation und damit zum Dickenzuwachs der SiO_2 -Schichten führt. Andernfalls kann eine derart homogene Belegung der Zwischenräume mit Siliziumdioxid nicht erklärt werden. Somit ist festzustellen, daß ein Reoxidationsschritt im weiteren Prozeßverlauf nach der Gatestrukturierung einen kritischen Prozeßschritt bei der

Verwendung von Titandioxid darstellt. Während dieses Prozeßschrittes erfolgt auch der Transfer in die Rutile-Phase.

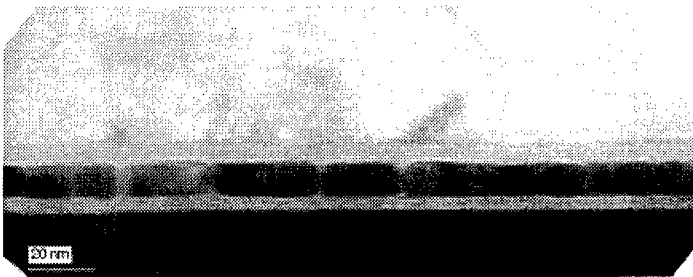


Abb. 6.15: In dieser TEM-Aufnahme in der Mitte zwischen Source und Drain eines Kurzkanaltransistors ist zu erkennen, daß Titandioxid große kristalline Körner ausbildet. Freie Bereiche in der Nähe der Korngrenzen werden durch Siliziumdioxid gefüllt.

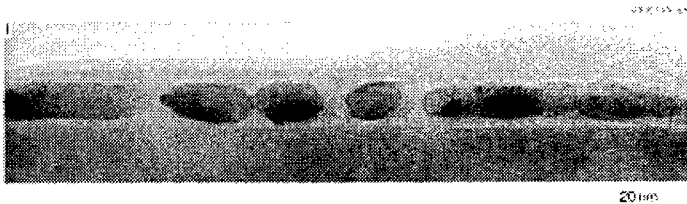


Abb. 6.16: TEM-Aufnahme einer sehr dünn präparierten Probe. Die Aufnahme zeigt einen Bereich in der Mitte einer großflächigen, planaren Kapazität.

6.4.5 Charakterisierung mittels Photoemissionsmikroskopie

Die Photoemissionsmikroskopie an ungestreßten planaren Kapazitäten zeigt eine hohe Flächendichte an Orten mit einem sehr hohen Leckstrom. In der Abbildung 6.17 ist die Photoemission einer planaren NMOS-Kapazität dargestellt. Jeder einzelne Punkt entspricht einem Leckstrompfad, durch den ein vergleichbar hoher Strom wie bei einem dielektrischen Durchbruch fließt. Die Aufnahme bestätigt die Erkenntnisse aus den Vergleichen der Gateleckströme verschiedener Kapazitäten und MOSFETs im Kapitel 6.4.3. und stimmt mit der strukturellen Analyse im Kapitel 6.4.4 überein. Die hohen Leckströme sind nicht auf die Eigenschaft des $\text{SiO}_2/\text{TiO}_2$ -Dielektrikums, sondern auf lokale Defekte zurückzuführen.

Zwei unmittelbar nacheinander folgende Analysen an einer planaren Kapazität zeigen, daß Orte mit einem hohen Leckstrom verschwinden oder neu hinzukommen. Ein Teil der

Strompfade ändert seine Emissionsintensität. Dies ist gleichbedeutend mit einer Veränderung des Leckstromes durch diesen Pfad. Zwischen den Analysen wurde kein weiterer elektrischer Streß angelegt und in beiden Analysen wurden die gleichen Meßparameter verwendet. Daraus ist zu schließen, daß Defekte und Haftstellen im Siliziumdioxid zwischen den TiO_2 -Körnern den hohen Strom verursachen. Die Umbesetzungen von Haftstellen beeinflussen den Elektronentransport durch diese Leckstrompfade.

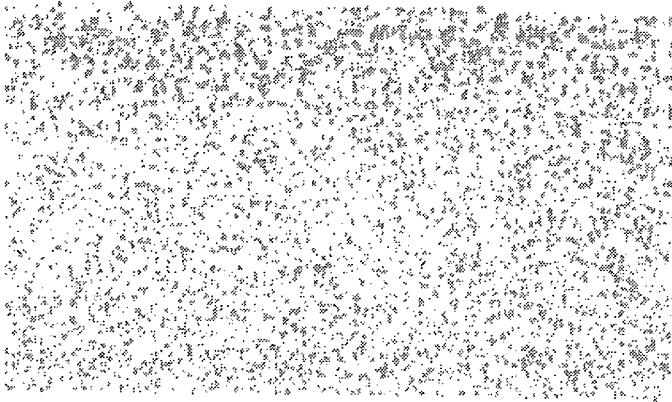


Abb. 6.17: Die Photoemissionsmikroskopie an einer planaren Kapazität zeigt eine hohe Dichte an Orten mit einem sehr hohen Leckstrom. Es ist nur die Photoemission dargestellt ohne einer photographischen Aufnahme der Teststruktur selbst.

6.4.6 Abschätzung für die Dielektrizitätskonstante

Die Dielektrizitätskonstante der Titandioxidschicht ist nicht unmittelbar bestimmbar, da der hohe Leckstrom Kapazitätsmessungen an planaren Kapazitäten nicht zuläßt. Mit Hilfe des Substrateffektes kann jedoch die effektive Oxiddicke des $\text{SiO}_2/\text{TiO}_2$ -Dielektrikums ermittelt werden. Damit erhält man eine Abschätzung für die Dielektrizitätskonstante der Titandioxidschicht. Die Substratdotierung N_{sub} kann ebenfalls mit Hilfe des Substrateffektes ermittelt werden, wenn MOSFETs mit einem SiO_2 -Gateoxid und derselben Substratdotierung zur Verfügung stehen. Die Gateoxiddicke ist dann bekannt und die Gateoxidkapazität kann in diesem Fall bestimmt werden. Der Substrateffekt beschreibt die Abhängigkeit der Einsatzspannung V_t von der Substratspannung V_{sub} [6.4]:

$$V_t(V_{\text{sub}}) = V_t(0) + \eta \cdot \left[\sqrt{2 \cdot \Psi_B + V_{\text{sub}}} - \sqrt{2 \cdot \Psi_B} \right] \quad (6.2)$$

$$\eta = \frac{\sqrt{2 \cdot \epsilon_{\text{DIE}} \cdot q \cdot N_{\text{sub}}}}{C_{\text{DIE}}} \quad (6.3)$$

$$2 \cdot \Psi_B = \frac{2 \cdot kT}{q} \ln \left(\frac{N_{\text{sub}}}{n_i} \right) \quad (6.4)$$

Ψ_B ist das Substratpotential, ϵ_{DIE} ist die effektive Dielektrizitätskonstante und C_{DIE} ist die Kapazität des Gatedielektrikums. n_i ist die intrinsische Ladungsträgerkonzentration von Silizium und beträgt $1,5 \cdot 10^{10} \text{ cm}^{-3}$ bei 27°C [6.4]. Der Faktor η wird jeweils graphisch ermittelt.

Mit Hilfe eines Vergleichswafers mit einem SiO_2 -Gateoxid wurde eine Substratdotierung von $N_{\text{sub}} = 6,8 \cdot 10^{17} \text{ cm}^{-3}$ bestimmt. Daraus und mit dem Faktor γ wurde die Kapazität des $\text{SiO}_2/\text{TiO}_2$ -Dielektrikums zu $C_{\text{DIE}} = 0,005 \text{ F/m}^2$ berechnet. Dies entspricht der Kapazität einer reinen SiO_2 -Schicht mit der Dicke von 6,9 nm. Die aus TEM-Aufnahmen ermittelte Summe aus Topoxidschicht und SiO_2 -Barrierschicht beträgt für Transistorstrukturen ca. 8 nm. Somit kann die TiO_2 -Schicht keinen wesentlichen Beitrag zur effektiven Oxiddicke liefern, auch wenn die Ungenauigkeit in der Bestimmung der Oxiddicken mittels TEM-Aufnahmen berücksichtigt wird. In der Literatur werden für TiO_2 relative Dielektrizitätskonstanten zwischen 4 und 86 genannt [6.11]. Treichel et al. ermittelten für TiO_2 -Schichten, die mittels PECVD abgeschieden und nicht weiter prozessiert wurden, eine relative Dielektrizitätskonstante von 22 bis 31 [6.15]. Guo et al. schätzten für die relative Dielektrizitätskonstante ihrer mittels JVD-Verfahren abgeschiedenen TiO_2 -Schichten ebenfalls einen Wert zwischen 20 und 30 ab [6.12].

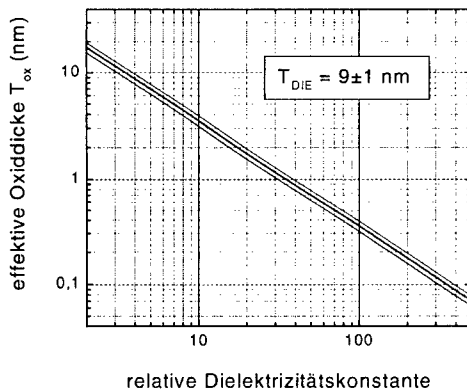


Abb. 6.18: Effektive SiO_2 -Dicke gegenüber der relativen Dielektrizitätskonstanten für ein Dielektrikum mit der physikalischen Schichtdicke von $9 \pm 1 \text{ nm}$.

In der Abbildung 6.18 ist für ein Dielektrikum mit physikalischer Dicke von 9 ± 1 nm die effektive Oxiddicke gegenüber der Dielektrizitätskonstanten aufgetragen. Die 9 nm entsprechen der Dicke der TiO_2 -Schicht, die aus TEM-Aufnahmen ermittelt wurde. Aus der Abbildung 6.18 ist zu entnehmen, daß eine relative Dielektrizitätskonstante im Bereich von 20 bis 30 eine effektive Oxiddicke von mehr als 1 nm ergibt. Für einen Wert 80 oder mehr wird der Beitrag zur effektiven Oxiddicke kleiner als 0,5 nm. Dies setzt noch einen Fehler in den obigen Messungen voraus, der nun allerdings um eine Größenordnung kleiner als die ermittelte effektive Oxiddicke ist. Unter Berücksichtigung des veröffentlichten Wertebereiches wird daher abgeschätzt, daß die TiO_2 -Schicht nach der vollständigen Prozessierung eine hohe relative Dielektrizitätskonstante im Bereich von 80 besitzen muß.

6.5 Zusammenfassung

Es wurde ein neues Verfahren zur Abscheidung von Titandioxid vorgestellt, das ausschließlich Standardprozesse verwendet. Das Verfahren ist geeignet, Titandioxid als Gatedielektrikum u.a. in verschiedene Prozeßkonzepte für Logiktechnologien zu implementieren. Augeranalysen weisen die Abscheidung von Titandioxid nach. Der Vergleich von physikalischen Schichtdicken und effektiver Oxiddicke ergab eine Abschätzung für die relative Dielektrizitätskonstante der Titandioxidschicht. Unter Berücksichtigung des veröffentlichten Wertebereiches wurde abgeschätzt, daß die TiO_2 -Schicht eine hohe relative Dielektrizitätskonstante im Bereich von 80 besitzen muß.

Einer der ersten Versuche, Titandioxid als Gatedielektrikum in einen kompletten CMOS-Prozeß bis zur ersten Metallisierungsebene zu implementieren, wurde charakterisiert. Um die Änderungen des Standardprozesses zu minimieren, wurde ein Basisprozeßkonzept verwendet. Dieses beinhaltet eine Barrierschicht aus Siliziumdioxid zwischen Siliziumsubstrat und Titandioxid und die Abscheidung eines Polysiliziumgates nach der Abscheidung von Titandioxid. Die vollständige Prozessierung erhöht die effektive Oxiddicke des gesamten Gatedielektrikums. Dies wird durch eine zusätzliche thermische Oxidation des Siliziumsubstrates und des Polysiliziums im Prozeßverlauf nach der Gatestrukturierung verursacht. Die zusätzliche Oxidation zeigt, daß das Basisprozeßkonzept durch eines der vorgestellten Prozeßkonzepte ersetzt werden muß, welches eine Barrierschicht verwendet, die eine Oxidation des Siliziumsubstrates verhindert.

Untersuchungen mittels Transmissionselektronenmikroskopie weisen nach, daß Titandioxid im verwendeten Gesamtprozeß große kristalline Körner bildet. Die Analyse der Leckstromdichte an verschiedenen Teststrukturen ergibt, daß durch strukturelle Defekte sehr hohe Leckströme verursacht werden können. Diese strukturellen Defekte äußern sich in weiten Bereichen zwischen Körnern aus Titandioxid, die mit Siliziumdioxid gefüllt sind. Eine Verteilung von Orten mit sehr hohen Leckströmen, die den strukturellen Defekten entsprechen, wurde mittels Photoemissionsmikroskopie nachgewiesen. Die Defekte treten bevorzugt bei großflächigen planaren Kapazitäten auf und werden durch die Möglichkeit zur mechanischen Streßrelaxation beeinflusst. Es wurden die Eingangskennlinien von NMOS-Kurzkanaltransistoren und PMOS-Kurzkanaltransistoren gezeigt. Die Transistoren besitzen

infolge der hohen effektiven Gateoxiddicke hohe Einsatzspannungen. Für PMOS-Kurzkanaltransistoren tritt ein Kink-Effekt auf, der durch eine zusätzliche laterale Ausdiffusion von Source und Drain verursacht wird. An den Isolationskanten entstehen dadurch parasitäre Kurzkanaltransistoren mit einer kleineren Einsatzspannung und einer kleineren Weite. Aus den Analysen ist insgesamt zu schließen, daß ein Reoxidationsprozeß nach der Gatestrukturierung die Verwendung von Titandioxid als Gatedielektrikum negativ beeinflußt. Der negative Effekt wird vermutlich verstärkt, wenn während dieses Reoxidationsprozesses ein Phasentransfer in die Rutile-Phase von Titandioxid stattfindet. Die physikalischen Analysen des Gatedielektrikums und die elektrische Charakterisierung von Kapazitäten und Transistoren zeigen, daß die vollständige Implementierung in einen CMOS-Prozeß möglich ist. Es werden jedoch wesentliche prozeßtechnische Änderungen erforderlich sein, um einen stabilen CMOS-Prozeß mit einem Gatedielektrikum aus Titandioxid oder einem anderen Metalloxid mit einer hohen Dielektrizitätskonstanten zu realisieren.

7 Anhang A

MOS-Strukturen und MOSFET

7.1 Das MOS-System – Energien und Potentiale

An dieser Stelle wird ein Überblick über die Bezeichnungen der verschiedenen Energien und Potentiale im MOS-System gegeben, welche in dieser Arbeit verwendet werden. Die Bezeichnungen sind auch im Verzeichnis der Symbole aufgelistet. Die Abbildung 7.1 zeigt folgende Energien und Potentiale: Energieniveau der Leitungsbandkante E_c , Energieniveau der Valenzbandkante E_v , Bandlücke des Halbleiters E_g , intrinsisches Energieniveau E_i , Fermi-niveau E_F , Substratpotential Ψ_B , Bandverbiegung an der Halbleiteroberfläche Ψ_s , Potentialbarriere zwischen Oxid und Halbleiter Φ_b , Potentialbarriere zwischen Oxid und Metallelektrode $\Phi_{b,m}$, Bandlücke des Oxides $E_{g,ox}$, Gatespannung V_g und Oxidspannung V_{ox} .

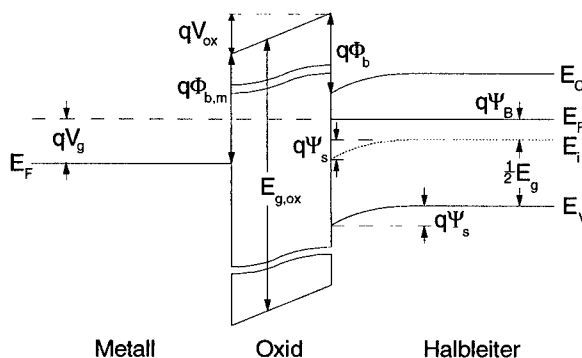


Abb. 7.1: Energiediagramm einer MOS-Struktur mit den Bezeichnungen wichtiger Energien und Potentiale.

7.2 Eingangskennlinien eines MOSFETs – wichtige Größen und Begriffe

Das Kapitel dient ebenfalls der Veranschaulichung gängiger Begriffe und der Definition von Symbolen, welche in den Kapiteln 2 und 5 verwendet werden. Die Abbildung 7.2 zeigt Eingangskennlinien eines MOSFETs. Es sind der Off-Strom I_{off} , der Drainstrom bei der Einsatzspannung des Transistors I_i und die Einsatzspannungen und Sättigungsdrainströme im linearen Bereich, $V_{t,\text{lin}}$ und $I_{d,\text{lin}}$, sowie im Sättigungsbereich, $V_{t,\text{sat}}$ und $I_{d,\text{sat}}$, eingezeichnet. Zudem sind die Unterschwellgerade und die Auswirkungen von „Drain-Induced Barrier Lowering“ (DIBL) und „Gate-Induced Drain Leakage“ (GIDL) dargestellt. „Gate-Induced Drain Leakage“ wird im Kapitel 7.5 dieses Anhangs kurz beschrieben. Das Kapitel 7.6 beschreibt die Ursache für Kurzkanaleffekte wie z.B. „Drain-Induced Barrier Lowering“.

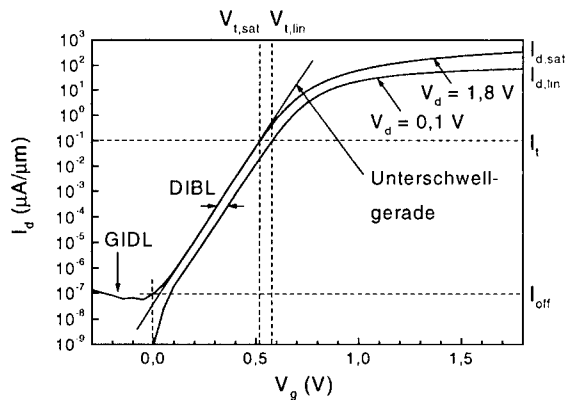


Abb. 7.2: Eingangskennlinien eines NMOSFETs im linearen Bereich ($V_d = 0,1 \text{ V}$) und im Sättigungsbereich ($V_d = 1,8 \text{ V}$). Es werden grundlegende Begriffe, Ströme und Einsatzspannungen erklärt.

7.3 Ladungsträgerbeweglichkeiten und effektives elektrisches Feld

In den Kapiteln 2 und 3 werden die Feldeffektbeweglichkeit und die effektive Beweglichkeit verwendet. Da in der Literatur oft nicht explizit auf den Unterschied verwiesen wird, soll dieser im folgenden dargelegt werden.

Die Feldeffektbeweglichkeit ist definiert als [7.1]:

$$\mu_{FE} = \frac{L}{W \cdot V_d \cdot C_{ox}} \cdot \frac{\partial I_d}{\partial V_g} \quad (7.1)$$

Dabei sind W und L die Weite und die Länge des Transistors, V_d die Drainspannung, V_g die Gatespannung, I_d der Drainstrom und C_{ox} die Oxidkapazität.

Mit der Transferleitfähigkeit g_m

$$g_m = \frac{\partial I_d}{\partial V_g} \quad (7.2)$$

gilt also folgender Zusammenhang:

$$\mu_{FE} = \frac{L}{W \cdot V_d \cdot C_{ox}} \cdot g_m \quad (7.3)$$

Für die effektive Beweglichkeit gilt:

$$\mu_{eff} = \frac{L}{W \cdot V_d \cdot C_{ox}} \cdot \frac{I_d}{V_g - V_t} = \frac{L}{W \cdot V_d} \cdot \frac{I_d}{|Q_t|} \quad (7.4)$$

Q_t ist die Inversionsladungsdichte und V_t die Einsatzspannung.

Die graphische Darstellung der effektiven Beweglichkeit gegenüber dem elektrischen Feld an der Substratoberfläche hängt von der Substratdotierung und der Substratspannung ab. Wird die effektive Beweglichkeit jedoch gegenüber dem effektiven elektrischen Feld E_{eff} aufgetragen, dann ergibt sich eine universelle Darstellung, die als „mobility universal relation“ bekannt ist [7.1]. E_{eff} wird folgendermaßen definiert:

$$E_{eff} = \frac{|Q_D + \eta \cdot Q_t|}{\epsilon_s} \quad (7.5)$$

Q_D ist die Ladungsdichte in der Verarmungszone und ϵ_s ist die Dielektrizitätskonstante des Halbleiters. Für Silizium ist also $\epsilon_s = 11,9 \cdot \epsilon_0$. Besitzt die Siliziumoberfläche eine $\langle 100 \rangle$ -Orientierung, so gilt für Elektronen $\eta = 1/2$ und für Löcher $\eta = 1/3$ [7.2].

7.4 „Buried Channel p-FET“

Zum besseren Verständnis des Kapitels 3.1 wird im folgenden der „Buried Channel p-FET“ beschrieben und der Unterschied zum „Surface Channel p-FET“ verdeutlicht. In der Abbildung 7.3 sind Simulationen der verschiedenen Dotierprofile für einen p-FET mit p^+ -Polysiliziumgate und mit n^+ -Polysiliziumgate gezeigt [7.3]. Die Einsatzspannung ist jeweils $V_t = -0,6$ V. Im letzteren Fall ist bei ca. $0,1 \mu m$ der p-n-Übergang zu erkennen. Für die weiteren schematischen Darstellungen werden rechteckige Profile und Ladungsverteilungen angenommen (siehe Abbildung 7.4). Zum besseren Verständnis wird zuerst der „Buried Channel p-FET“ vom Typ „normally-on“ diskutiert. Der Bereich des vergrabenen Kanals ist

in der Abbildung 7.5 mit x_{bc} bezeichnet. Bei einer Gatespannung $V_g = 0\text{ V}$ existieren im Substrat zwei Verarmungszonen, an der Oberfläche des p-dotierten Gebietes und im p-n-Übergang. Die Weiten dieser Verarmungszonen sind in der Abbildung 7.5a mit x_1 und $w_1 + w_2$ bezeichnet.

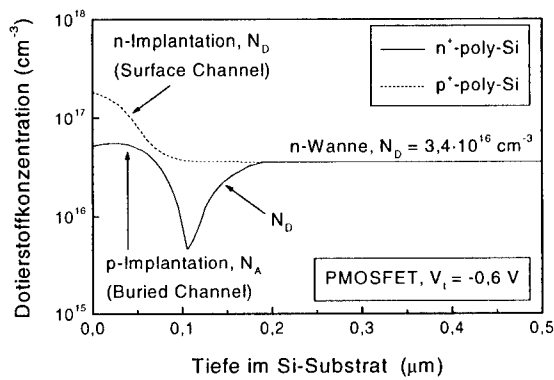


Abb. 7.3: Simulationen der Dotierstoffprofile für je einen p-FET mit n^+ -Polysiliziumgate und p^+ -Polysiliziumgate und $V_t = -0,6\text{ V}$ [7.3].

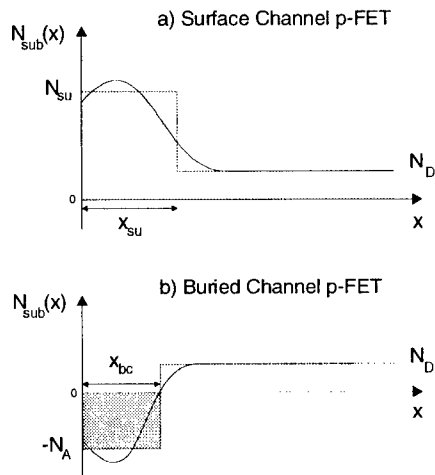


Abb. 7.4: Schematische Darstellung der Dotierstoffprofile für „Surface Channel p-FET“ und „Buried Channel p-FET“.

Das Banddiagramm ist in der Abbildung 7.5b skizziert. In diesem Zustand existiert ein p-leitender Bereich zwischen Source und Drain. Mit positiveren Gatespannungen wird die Weite x_1 größer, bis das p-dotierte Gebiet (Akzeptoren, N_A) völlig verarmt ist. Dies ist in den Abbildungen 7.5c und 7.5d dargestellt. Die völlige Verarmung ereignet sich zuerst in Drainnähe und wird analog zum Oberflächenkanaltransistor als „pinch-off“ bezeichnet [7.4]. Die beteiligten Kapazitäten sind in der Abbildung 7.5e gezeigt (Abbildung gemäß [7.4]). Die Kapazität des vollständig verarmten „Buried Channel“-Gebietes ist: $C_{bc} = \epsilon_{Si} \cdot A / x_{bc}$.

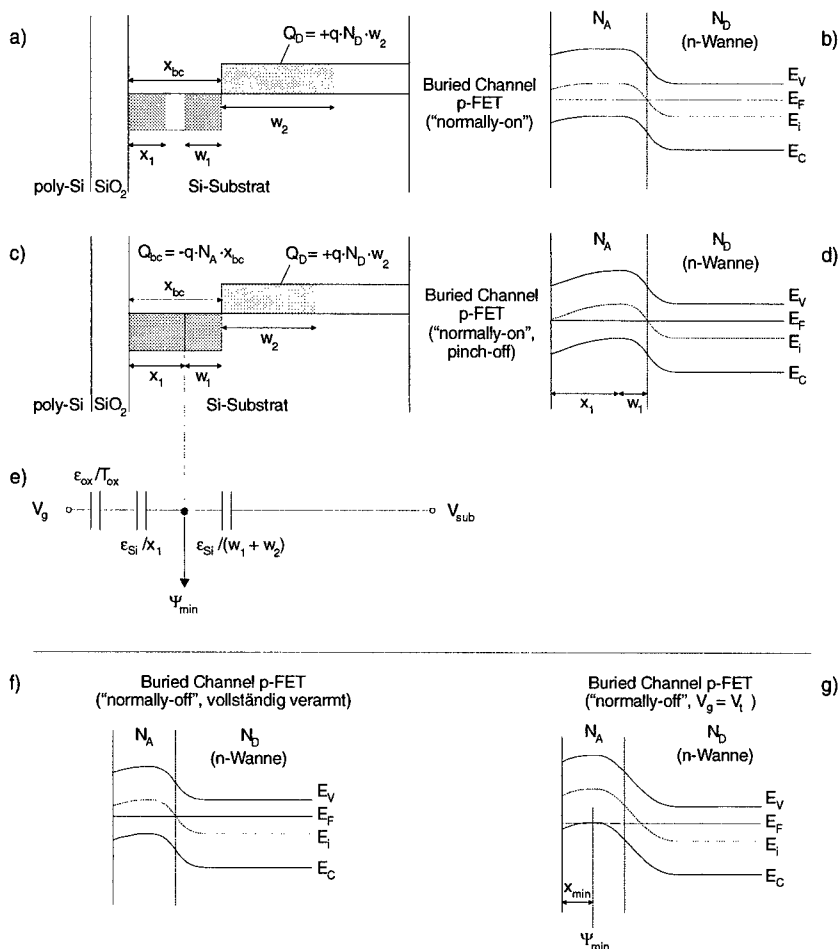


Abb. 7.5: Schematische Darstellungen zur Erklärung eines „Buried Channel p-FET“.

Über die Dotierstoffkonzentration N_A oder die Tiefe des Implantationsprofils läßt sich erreichen, daß das „Buried Channel“-Gebiet bei $V_g = 0$ über die gesamte Transistorlänge völlig verarmt ist. Der „Buried Channel p-FET“ ist dann vom Typ „normally-off“. Dieser Transistortyp wurde in den CMOS-Technologien mit einem n^+ -Polysiliziumgate verwendet. Das Banddiagramm ist in der Abbildung 7.5f skizziert. Bei $V_g = V_t < 0$ ergibt sich um den Ort des minimalen Potentials x_{\min} ein p-leitendes Gebiet (siehe Abbildung 7.5g). Dies ist analog zu dem Fall, der in der Abbildung 7.5a skizziert ist. Für $V_g < V_t$ bewirkt die weitere Bandverbiegung an der Substratoberfläche, daß sich das p-leitende Gebiet bis zur Grenzfläche zum Gateoxid erstreckt. Das bedeutet, daß der „Buried Channel p-FET“ zunächst in einer Tiefe x_{\min} leitfähig wird, und sich das Kanalgebiet dann bei vollständiger Ansteuerung bis zur Oberfläche hin erstreckt.

Eine Gegenüberstellung von Ladungen, elektrischem Feld und Potential für den „Surface Channel p-FET“ und den „Buried Channel p-FET“ enthält die Abbildung 7.6. Die Abbildung 7.7 zeigt Simulationen der maximalen Potentialbarriere zwischen Source und Drain in Abhängigkeit von der Substrattiefe [7.3]. Für den p-FET mit n^+ -Polysiliziumgate ist hier zu erkennen, daß die kleinste Potentialbarriere ca. 50 nm tief im Substrat liegt und dort den Beginn der Kanalleitfähigkeit bestimmt. Dies ist charakteristisch für einen „Buried Channel-FET“.

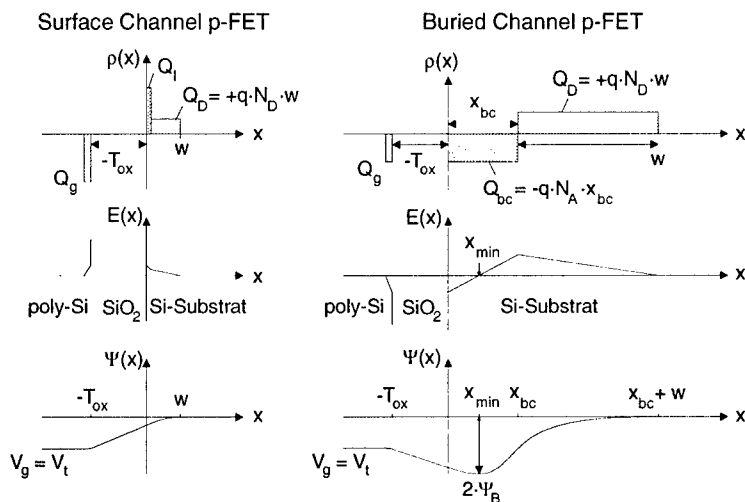


Abb. 7.6: Ladungen, elektrisches Feld und Potential für den „Surface Channel p-FET“ und den „Buried Channel p-FET“.

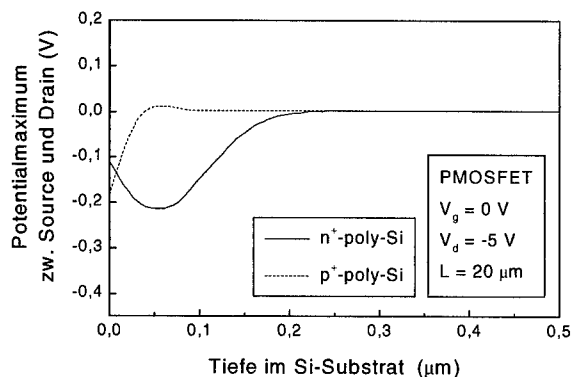


Abb. 7.7: Simulationen der maximalen Potentialbarriere zwischen Source und Drain in Abhängigkeit von der Substrattiefe [7.3]. Die p-FET mit n^+ -Polysiliziumgate und mit p^+ -Polysiliziumgate haben die gleiche Einsatzspannung, ähnlich wie bei der Simulation in der Abbildung 7.3.

7.5 „Gate-Induced Drain Leakage“

Der „Gate-Induced Drain Leakage“-Strom ist ein Tunnelstrom zwischen dem Valenzband und dem Leitungsband im Silizium. Im Kapitel 5 wird gezeigt, daß „Gate-Induced Drain Leakage“ (GIDL) eine wichtige Rolle bei der Degradation eines NMOSFETs durch einen „Soft Breakdown“ im Gateoxid spielt. Daher wird „Gate-Induced Drain Leakage“ im folgenden anhand eines NMOSFETs erklärt. Für den PMOSFET ist der Sachverhalt ähnlich, nur mit umgekehrten Vorzeichen für die Potentialdifferenzen und umgekehrten Beiträgen von Elektronen und Löchern.

Im NMOSFET sind die Diffusionsgebiete n-dotiert und die Wanne p-dotiert. Bei einer negativen Potentialdifferenz zwischen Gate und Drain wird das n-dotierte Gebiet im Überlappbereich zwischen Gate und Drain an der Substratoberfläche invertiert. Dies ist z.B. der Fall, wenn am Drain die Versorgungsspannung V_{DD} anliegt und die Gatespannung sehr klein bzw. negativ wird. Das invertierte n-Gebiet hat einen Anschluß an die p-dotierte Wanne, die sich bei einer entsprechenden Gatespannung an der Substratoberfläche in Akkumulation befindet. In der Raumladungszone des invertierten n-Gebietes am Drainkontakt werden Elektron-Loch-Paare generiert. Die Elektronen tunneln nun vom Valenzband in das Leitungsband und liefern einen Beitrag zum Strom am Drainkontakt. Die Löcher gelangen in die p-dotierte Wanne und ergeben einen Beitrag zum Strom am Wannenkontakt. Dessen Betrag ist gleich dem Betrag des Stromes an Elektronen am Drainkontakt.

Die Bandverbiegung ist abhängig vom elektrischen Feld an der Siliziumsubstratoberfläche. Die Potentialbarriere, welche die Elektronen durchtunneln müssen, entspricht näherungsweise einem Dreieckspotential. Damit weist die GIDL-Stromdichte eine entsprechende Abhängigkeit vom elektrischen Feld auf wie die Fowler-Nordheim-Tunnelstromdichte, die zu Beginn des Kapitels 4 angegeben ist (siehe z.B. [7.1]).

In der Abbildung 7.1 ist für die Eingangskennlinie bei $V_D = V_{DD} = 1,8 \text{ V}$ zu erkennen, daß der GIDL bei $V_g \approx 0,1 \text{ V}$ größer als der Drainstrom des Transistors wird und für $V_g < 0$ weiter ansteigt.

7.6 Kurzkanaleffekte

Im Kapitel 6 werden zur Erklärung der Eingangskennlinien von PMOSFETs mit einer Gatelänge von $0,20 \mu\text{m}$ Kurzkanaleffekte herangezogen. Die Ursache für Kurzkanaleffekte wird daher kurz dargelegt.

Wenn die Weiten der Raumladungszonen von Source und Drain in der Größenordnung der Kanallänge liegen, werden die Potentialverhältnisse im Inversionskanal zusätzlich von der lateralen Potentialverteilung zwischen Source und Drain abhängig. Damit ergibt sich mit dem vertikalen elektrischen Feld eine zweidimensionale Potentialabhängigkeit. Diese Zweidimensionalität bewirkt unterschiedliche Effekte, die allgemein als Kurzkanaleffekte („short channel effects“) bezeichnet werden. Zwei wichtige Kurzkanaleffekte sind die Abnahme der Einsatzspannung mit der Kanallänge („ V_t roll-off“) und die Absenkung der maximalen Potentialbarriere zwischen Source und Drain („Drain-Induced Barrier Lowering“). „Drain-Induced Barrier Lowering“ (DIBL) äußert sich in einer Abhängigkeit der Eingangskennlinie im Unterschwellbereich von der Drainspannung und ist in der Abbildung 7.2 zu erkennen. Eine Beschreibung der Kurzkanaleffekte erfolgt mit Hilfe des Modells der Ladungsteilung („charge sharing model“). Dieses berücksichtigt, daß die Raumladung im Substrat unterhalb des Gates teilweise der Oberflächenraumladungszone sowie den Raumladungszonen der p-n-Übergänge zuzuordnen ist. Eine kompakte Beschreibung dieses Modells und der Kurzkanaleffekte findet sich z.B. in [7.1].

8 Anhang B

Statistik

Bei der Charakterisierung der Zuverlässigkeit von Oxiden werden für die Auswertung der statistisch verteilten Durchbruchereignisse die Weibull-Verteilung und die Poisson-Verteilung verwendet. Im folgenden werden daher die allgemeinen Formulierungen für diese beiden Verteilungen angegeben.

8.1 Allgemeines zur Statistik

Ist $f(t)$ die Wahrscheinlichkeitsdichte für einen Fehler zur Zeit t , so gilt für die Verteilungsfunktion $F(t)$:

$$F(t) = \int_{-\infty}^t f(t') dt' \quad (8.1)$$

Die momentane Fehlerrate $\lambda(t)$ ist definiert als:

$$\lambda(t) = \frac{f(t)}{1 - F(t)} \quad (8.2)$$

Die Größe $\lambda(t)$ gibt die Fehlerrate der verbliebenen fehlerfreien Proben zur Zeit t für die unmittelbar folgende Zeit an.

8.2 Die Weibull-Verteilung

Im folgenden wird die allgemeine Form der Weibullverteilung beschrieben [8.1].

Die stetige Zufallsgröße X besitzt eine dreiparametrische Weibullverteilung mit den Parametern a , b und c ($a > 0$, $b > 0$, c beliebig reell), wenn ihre Dichte durch

$$f(x) = 0 \quad \text{für } x \leq c \quad (8.3)$$

$$f(x) = \frac{b}{a} \cdot \left[\frac{x-c}{a} \right]^{b-1} e^{-\left(\frac{x-c}{a}\right)^b} \quad \text{für } x > c \quad (8.4)$$

und ihre Verteilungsfunktion durch

$$F(x) = 0 \quad \text{für } x \leq c \quad (8.5)$$

$$F(x) = 1 - e^{-\left(\frac{x-c}{a}\right)^b} \quad \text{für } x > c \quad (8.6)$$

gegeben ist.

Es sind a der Maßstabsparameter, b der Formparameter und c der Lageparameter der Weibull-Verteilung. Ist $a = 1$, so liegt die sogenannte reduzierte Weibull-Verteilung vor. Für $c = 0$ erhält man die zweiparametrische Weibull-Verteilung. Sind $a = 1/\alpha$, $b = 1$ und $c = 0$, so ergibt sich als Spezialfall der Weibull-Verteilung die Exponentialverteilung.

Die Weibull-Verteilung gehört zu den sog. Extremwertverteilungen, denn sie ergibt sich als Grenzverteilung für das Minimum einer großen Zahl unabhängiger Zufallsgrößen [8.2]. Das rechtfertigt z.B. ihre Verwendung bei der Untersuchung der Lebensdauer eines Systems aus hintereinander geschalteten Elementen, die durch die Lebensdauer des schwächsten Gliedes in der Elementenkette charakterisiert werden kann. So findet sie vor allem bei Lebensdauer- und Zuverlässigkeitsanalysen, z.B. als Modell zur Beschreibung von Alterungs- und Verschleißvorgängen, Anwendung [8.1].

Der Erwartungswert und die Streuung der Weibull-Verteilung sind:

$$\mu = E(X) = c + a \cdot \Gamma\left(1 + \frac{1}{b}\right) \quad (8.7)$$

$$\sigma^2 = \text{Var}(X) = a^2 \left[\Gamma\left(1 + \frac{2}{b}\right) - \Gamma^2\left(1 + \frac{1}{b}\right) \right] \quad (8.8)$$

Γ bezeichnet die Gammafunktion mit den Eigenschaften:

$$\begin{aligned} \Gamma(1) &= 1 \\ \Gamma(1/2) &= \sqrt{\pi} \\ \Gamma(x) &= (x-1) \cdot \Gamma(x-1) \end{aligned} \quad (8.9)$$

8.3 Die Poisson-Verteilung

Im folgenden wird die allgemeine Formulierung der Poisson-Verteilung angegeben [8.1].

Die diskrete Zufallsgröße X , welche die Werte $k = 0, 1, 2, \dots$ annehmen kann mit den Wahrscheinlichkeiten

$$p_k = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad (8.10)$$

besitzt eine Poisson-Verteilung mit dem Parameter λ ($\lambda > 0$).

Der Erwartungswert und die Streuung sind:

$$\begin{aligned}\mu &= E(X) = \lambda \\ \sigma^2 &= \text{Var}(X) = \lambda\end{aligned}\quad (8.11)$$

Es gilt außerdem:

$$\begin{aligned}\sum_{k=0}^{\infty} p_k &= 1 \\ p_{k+1} &= \frac{\lambda}{k+1} \cdot p_k\end{aligned}\quad (8.12)$$

Gemäß der Aussage des Grenzwertsatzes von Poisson geht die Poisson-Verteilung aus der Binominalverteilung hervor, wenn dort n gegen unendlich und p gegen Null streben, und zwar so daß $n \cdot p = \lambda$ konstant bleibt.

$$\lim_{\substack{n \rightarrow \infty \\ np = \lambda}} \binom{n}{k} p^k (1-p)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda} \quad (k = 0, 1, 2, \dots) \quad (8.13)$$

Daher nennt man die Poisson-Verteilung auch die Verteilung der seltenen Ereignisse.

Defektdichte und Ausbeute

Jeder fertig prozessierte Wafer ist mit einer Anzahl von Defekten belastet, die zu fehlerhaften elektrischen Schaltungen führen können. Die Ausbeute an funktionierenden Chips wird als „Yield“ bezeichnet. Die Chipfläche sei A , und die Defekte seien zufällig auf dem Wafer mit einer Defektdichte D verteilt. Das Produkt $A \cdot D$ ergibt dann die mittlere Anzahl von Defekten auf einem Chip. Die Wahrscheinlichkeit, insgesamt m Defekte auf einem Chip zu finden, ist durch die Poisson-Verteilung gegeben:

$$P(m) = \frac{(A \cdot D)^m}{m!} \cdot e^{-A \cdot D} \quad (8.14)$$

„Yield“ Y ist die Wahrscheinlichkeit, einen Chip ohne Defekt zu finden. Für diesen Fall ist $m = 0$ und es gilt:

$$Y = P(0) = e^{-A \cdot D} = 1 - F \quad (8.15)$$

Diese Gleichung ist auch für das analoge Problem der Verteilung von Oxiddefekten auf einer aktiven Fläche A gültig. Die Voraussetzung ist, daß die Defekte zufällig auf der Fläche A verteilt sind.

Literaturverzeichnis

Kapitel 1

- 1.1 J. Bardeen, W. H. Brattain, The Transistor, A Semiconductor Triode, *Phys. Rev.* **74**, S. 230 (1948)
- 1.2 W. H. Brattain, J. Bardeen, Nature of the Forward Current in Germanium Point Contacts, *Phys. Rev.* **74**, S. 231 (1948)
- 1.3 W. Shockley, The Theory of p-n Junctions in Semiconductors and p-n Junction Transistors, *Bell System Technical Journal*, S. 435 (1949)
- 1.4 Deutsches Patent, Nr. 814.487
- 1.5 C. Y. Chang, S. M. Sze, *ULSI Technology*, Verlag: The McGraw-Hill Companies, ISBN 0-07-063062-3 (1996)
- 1.6 Semiconductor Industry Association, *International Technology Roadmap for Semiconductors* (1998)

Kapitel 2

- 2.1 G. Barbottin, A. Vapaille, *Instabilities in Silicon Devices, Silicon Passivation and Related Issues, Volume 2*, Verlag: Elsevier Science Publishers, ISBN 0-444-70016-1 (1989)
- 2.2 S. M. Sze, *Physics of Semiconductor Devices*, Verlag: John Wiley & Sons, ISBN 0-471-05661-8 (1981)
- 2.3 E. H. Nicollian, J. R. Brews, *MOS (Metal Oxide Semiconductor) Physics and Technology*, Verlag: John Wiley & Sons, ISBN 0-471-08500-6 (1982)
- 2.4 R. F. Pierret, *Semiconductor Device Fundamentals*, Verlag: Addison-Wesley Publishing Company, ISBN 0-201-54393-1 (1996)
- 2.5 F. Stern, W. E. Howard, Properties of Semiconductor Surface Inversion Layers in the Electric Quantum Limit, *Phys. Rev.* **163**(3), S. 816 (1967)
- 2.6 T. Ando, A. B. Fowler, F. Stern, Electronic properties of two-dimensional systems, *Reviews of Modern Physics* **54**(2), S. 437-672 (1982)
- 2.7 W. Liu, X. Jin, Y. King, C. Hu, An Efficient and Accurate Compact Model for Thin-Oxide-MOSFET Intrinsic Capacitance Considering the Finite Charge Layer Thickness, *IEEE Trans. on Electron Devices* **46**(5), S. 1070 (1999)
- 2.8 H. Reisinger, H. Oppolzer, W. Hönlein, Thickness Determination of Thin SiO₂ on Silicon, *Solid-State Electronics* **35**(6), S. 797 (1992)
- 2.9 B. Riccò, P. Olivo, T. N. Nguyen, T.-S. Kuan, G. Ferriani, Oxide-Thickness Determination in Thin-Insulator MOS Structures, *IEEE Trans. on Electron Devices* **35**(4), S. 432 (1988)
- 2.10 R. Rios, N. D. Arora, Determination of Ultra-Thin Gate Oxide Thickness for CMOS Structures Using Quantum Effects, *Technical Digest of the 1994 International Electron Device Meeting (IEDM)*, S. 613 (1994)

- 2.11 C. Bowen, C. L. Fernando, G. Klimeck, A. Chatterjee, D. Blanks, R. Lake, J. Hu, J. Davis, M. Kulkarni, S. Hattangady, I. Chen, Physical Oxide Thickness Extraction and Verification using Quantum Mechanical Simulation, *Technical Digest of the 1997 International Electron Device Meeting (IEDM)*, S. 869 (1997)
- 2.12 K. Yang, Y.-C. King, C. Hu, Quantum Effect in Oxide Thickness Determination From Capacitance Measurements, *Proceedings of the 1999 Symposium on VLSI Technology* (1999)
- 2.13 C.-H. Choi et al., C-V and Gate Tunneling Current Characterization of Ultra-Thin Gate Oxide MOS ($t_{ox}=1.3-1.8\text{nm}$), *Proceedings of the 1999 Symposium on VLSI Technology* (1999)
- 2.14 J. S. Brugler, P. G. A. Jespers, Charge Pumping in MOS Devices, *IEEE Trans. on Electron Devices*, ED-16(3), S. 297 (1969)
- 2.15 G. Groeseneken, H. E. Maes, N. Beltrán, R. F. DeKeersmaecker, A Reliable Approach to Charge-Pumping Measurements in MOS Transistors, *IEEE Trans. on Electron Devices* ED-31(1), S. 42 (1984)
- 2.16 T. Hori, *Gate Dielectrics and MOS ULSIs*, Verlag: Springer, ISBN 3-540-63182-8 (1997)
- 2.17 A. G. Sabnis, J. T. Clemens, Characterization of the electron mobility in the inverted <100> Si surface, *Technical Digest of the 1979 International Electron Device Meeting (IEDM)*, S. 18 (1979)
- 2.18 S. Takagi, A. Toriumi, M. Iwase, H. Tango, On the Universality of Inversion Layer Mobility in Si MOSFET's: Part I – Effects of Substrate Impurity Concentration, *IEEE Trans. on Electron Devices* 41(12), S. 2357 (1994)
- 2.19 E. Takeda, C. Y. Yang, A. Miura-Hamada, *Hot Carrier Effects in MOS Devices*, Verlag: Academic Press, ISBN 0-12-682240-9 (1995)
- 2.20 G. Van den Bosch, G. Groeseneken, H. E. Maes, Critical analysis of the substrate hot-hole injection technique, *Solid-State Electronics* 37(3), S. 393 (1994)
- 2.21 W. Shockley, Problems related to p-n Junctions in silicon, *Solid-State Electronics* 2(35), S. 35 (1961)
- 2.22 P. E. Cottrell, R. R. Troutman, T. H. Ning, Hot-Electron Emission in n-Channel IGFETs, *IEEE Trans. on Electron Devices* 26, S. 520 (1979)
- 2.23 C. Hu, Lucky-electron model for hot-electron emission, *Technical Digest of the 1979 International Electron Device Meeting (IEDM)*, S. 22 (1979)
- 2.24 A. Schwerin, W. Hänsch, W. Weber, The relationship between oxide charge and device degradation: A comparative study of n- and p-MOSFET's, *IEEE Trans. on Electron Devices* ED-34(12), S. 2493 (1987)
- 2.25 P. Heremans, R. Bellens, G. Groeseneken, H. E. Maes, Consistent Model for the Hot-Carrier Degradation in n-Channel and p-Channel MOSFET's, *IEEE Trans. on Electron Devices* 35(12), S. 2194 (1988)
- 2.26 C. Hu, S. C. Tam, F.-C. Hsu, P.-K. Ko, T.-Y. Chan, K. W. Terrill, Hot-Electron-Induced MOSFET Degradation – Model, Monitor, and Improvement, *IEEE Trans. on Electron Devices* ED-32(2), S. 375 (1985)
- 2.27 Q. Wang, M. Brox, H. Krautschneider, W. Weber, Explanation and Model for the Logarithmic Time Dependence of p-MOSFET Degradation, *IEEE Electron Device Letters* 12(5), S. 218 (1991)
- 2.28 M. Brox, A. v. Schwerin, Q. Wang, W. Weber, A Model for the Time- and Bias-Dependence of p-MOSFET Degradation, *IEEE Trans. on Electron Devices* 41(7), S. 1184 (1994)
- 2.29 R. Thewes, *Degradation analoger CMOS-Schaltungen durch heiße Ladungsträger*, Dissertation (1995)
- 2.30 K. O. Jeppson, C. M. Svensson, Negative bias stress of MOS devices at high electric fields and degradation of MNOS devices, *Journal of Applied Physics* 48(5), S. 2004 (1977)
- 2.31 J. F. Zhang, W. Eccleston, Positive Bias Temperature Instability in MOSFET's, *IEEE Trans. on Electron Devices* 45(1), S. 116 (1998)
- 2.32 K. Uwasawa, T. Yamamoto, T. Mogami, A New Degradation Mode of Scaled p⁺ Polysilicon Gate pMOSFETs Induced by Bias Temperature (BT) Instability, *Technical Digest of the 1995 International Electron Device Meeting (IEDM)*, S. 871 (1995)

- 2.33 Y. Y. Chen, D. L. Kwong, M. Gardner, J. Fulford, Study of NBTI Stress in P⁺-Poly PMOSFETs with Gate Oxides Grown on Nitrogen Implanted Si Substrates. *Proceedings of the 1998 European Solid State Device Research Conference (ESSDERC)*, S. 292 (1998)
- 2.34 S. Ogawa, M. Shimaya, N. Shiono, Impact of Negative-Bias Temperature Instability on the Lifetime of Single-Gate CMOS Structures with Ultrathin (4-6 nm) Gate Oxides. *Japanese Journal of Applied Physics* 35, S. 1484 (1996)
- 2.35 T. Yamamoto, K. Uwasawa, T. Mogami, Bias Temperature Instability in Scaled p⁺ Polysilicon Gate p-MOSFET's, *IEEE Trans. on Electron Devices* 46(5), S. 921 (1999)
- 2.36 C. Schlönder, R. Brederlow, P. Wiczorek, Claus Dahl, Jürgen Holz, Michael Röhner, Sylvia Kessel, Volker Herold, Karl Goser, Werner Weber, Roland Thews, Trapping Mechanisms in Negative Bias Temperature Stressed p-MOSFETs. *Proceedings of the 1999 European Symposium on Reliability of Electron Devices, Failure Physics and Analysis (ESREF)* (1999)
- 2.37 C. R. Barrett, R. C. Smith, Failure modes and reliability of dynamic rams. *Technical Digest of the 1976 International Electron Device Meeting (IEDM)* (1976)
- 2.38 E. Anolick, G. Nelson, Low field time dependent dielectric integrity. *Proceedings of the 1979 International Reliability Physics Symposium (IRPS)*, S. 8 (1979)
- 2.39 D. L. Crook, Method of determining reliability screens for time dependent dielectric breakdown. *Proceedings of the 1979 International Reliability Physics Symposium (IRPS)*, S. 1 (1979)
- 2.40 A. Berman, Time-Zero Dielectric Reliability Test by a Ramp Method *Proceedings of the 1981 International Reliability Physics Symposium (IRPS)*, S. 204 (1981)
- 2.41 J. McPherson, D. A. Baglee, Acceleration factors for thin gate oxide stressing. *Proceedings of the 1985 International Reliability Physics Symposium (IRPS)*, S. 1 (1985)
- 2.42 J. McPherson, D. A. Baglee, Acceleration Factors for Thin Oxide Breakdown", *Journal of the Electrochemical Society* 132(8), S. 1903 (1985)
- 2.43 J. McPherson, V. Reddy, H. Mogul, Field-enhanced Si-Si bond-breaking mechanism for time-dependent dielectric breakdown in thin-film SiO₂ dielectrics. *Applied Physics Letters* 71(8), S. 1101 (1997)
- 2.44 J. McPherson, H. Mogul, Impact of mixing of disturbed bonding states on time-dependent dielectric breakdown in SiO₂ thin films. *Applied Physics Letters*, 71, S. 3721 (1997)
- 2.45 J. McPherson, H. Mogul, Disturbed bonding states in SiO₂ thin-films and their impact on time-dependent dielectric breakdown, *Proceedings of the 1998 International Reliability Physics Symposium (IRPS)*, S. 47 (1998)
- 2.46 J. McPherson, H. Mogul, Underlying physics of the thermochemical E model in describing low-field time-dependent dielectric breakdown in SiO₂ thin films. *Journal of Applied Physics* 84(3), S. 1513 (1998).
- 2.47 M. Kimura, Field and Temperature Acceleration Model for Time-Dependent Dielectric Breakdown. *IEEE Trans. on Electron Devices* 46(1), S. 220 (1999).
- 2.48 I. C. Chen, S. E. Holland, C. Hu, Electrical Breakdown in Thin Gate and Tunneling Oxides. *IEEE Trans. on Electron Devices* ED-32(2), S. 413 (1985)
- 2.49 I. C. Chen, S. E. Holland, C. Hu, A quantitative physical model for time-dependent breakdown in SiO₂. *Proceedings of the 1985 International Reliability Physics Symposium (IRPS)*, S. 24 (1985)
- 2.50 K. F. Schuegraf, C. Hu, Hole injection oxide breakdown model for very low voltage lifetime extrapolation", *Proceedings of the 1993 International Reliability Physics Symposium (IRPS)*, S. 7 (1993)
- 2.51 K. F. Schuegraf, C. Hu, Hole Injection SiO₂ Breakdown Model for Very Low Voltage Lifetime Extrapolation. *IEEE Trans. on Electron Devices* 41(5), S. 761 (1994)
- 2.52 K. F. Schuegraf, C. Hu, Reliability of thin SiO₂. *Semiconductor Science and Technology* 9, S. 989 (1994)
- 2.53 J. C. Lee, I. C. Chen, C. Hu, Modeling and Characterization of Gate Oxide Reliability. *IEEE Trans. on Electron Devices* 35(12), S. 2268 (1988)

- 2.54 J. S. Suehle, P. Chaparala, C. Messick, W. M. Miller, K. C. Boyko, Field and Temperature Acceleration of Time-Dependent Dielectric Breakdown in Intrinsic Thin SiO₂, *Proceedings of the 1994 International Reliability Physics Symposium (IRPS)*, S. 120 (1994)
- 2.55 N. Shiono, M. Itsumi, A Lifetime Projection Method Using Series Model and Acceleration Factors for TDDB failures of Thin Gate Oxides, *Proceedings of the 1993 International Reliability Physics Symposium (IRPS)*, S. 1 (1993)
- 2.56 M. Kimura, Oxide Breakdown Mechanism and Quantum Physical Chemistry for Time-Dependent Dielectric Breakdown, *Proceedings of the 1997 International Reliability Physics Symposium (IRPS)*, S. 190 (1997)
- 2.57 J. W. McPherson, V. Reddy, K. Banerjee, H. Le, Comparison of E and 1/E TDDB Models for SiO₂ under long-term/low-field test conditions, *Technical Digest of the 1998 International Electron Device Meeting (IEDM)* (1998)
- 2.58 A. Teramoto, H. Umeda, K. Azamawari, K. Kobayashi, K. Shiga, J. Komori, Y. Ohno, H. Miyoshi, Study of Oxide Breakdown under Very Low Electric Field, *Proceedings of the 1999 International Reliability Physics Symposium (IRPS)*, S. 66 (1999)
- 2.59 C. Hu, Q. Lu, A Unified Gate Oxide Reliability Model, *Proceedings of the 1999 International Reliability Physics Symposium (IRPS)*, S. 47 (1999)
- 2.60 K. P. Cheung, A physics-based, unified gate-oxide breakdown model, *Technical Digest of the 1999 International Electron Device Meeting (IEDM)* (1999)
- 2.61 R. Storm, *Wahrscheinlichkeitsrechnung, mathematische Statistik und statistische Qualitätskontrolle*, Verlag: Fachbuchverlag Leipzig, ISBN 3-343-00871-0 (1995)
- 2.62 E. Y. Wu, W. W. Abadeer, L. Han, S. Lo, G. R. Hueckel, Challenges for Accurate Reliability Projections in the Ultra-thin Oxide Regime, *Proceedings of the 1999 International Reliability Physics Symposium (IRPS)*, S. 57 (1999)

Kapitel 3

- 3.1 G. J. Hu, Design Tradeoffs Between Surface and Buried-Channel FET's, *IEEE Trans. on Electron Devices ED-32*(3), S. 584 (1985)
- 3.2 T. Hori, *Gate Dielectrics and MOS ULSIs*, Verlag: Springer, ISBN 3-540-63182-8 (1997)
- 3.3 R. B. Fair, Physical Models of Boron Diffusion in Ultrathin Gate Oxides, *Journal of the Electrochemical Society* 144(2), S. 708 (1997)
- 3.4 P. J. Tobin, Y. Okada, S. A. Ajuria, V. Lakhotia, W. A. Feil, R. I. Hedge, Furnace formation of silicon oxynitride thin dielectrics in nitrous oxide (N₂O): the role of nitric oxide (NO), *Journal of Applied Physics* 75(3), S. 1811 (1994)
- 3.5 Y. Okada, P. J. Tobin, K. G. Reid, R. I. Hegde, B. Maiti, S. A. Ajuria, Gate Oxynitride Grown in Nitric Oxide (NO), *Proceedings of the 1994 Symposium on VLSI Technology*, S. 105 (1994)
- 3.6 A. R. Mattheus, *Charakterisierung von dielektrischen Schichten für EEPROM-Zellen*, Dissertation (1998)
- 3.7 H. R. Soleimani, B. S. Doyle, A. Philipossian, Formation of Ultrathin Nitrided Oxides by Direct Nitrogen Implantation into Silicon, *Journal of the Electrochemical Society* 142(8), S. L132 (1995)
- 3.8 T. Kuroi, T. Yamaguchi, M. Shirahata, Y. Okumura, Y. Kawasaki, M. Inuishi, N. Tsubouchi, Novel NICE (nitrogen implantation into CMOS gate electrode and source-drain) structure for high reliability and high performance 0.25 μ m dual gate CMOS, *Technical Digest of the 1993 International Electron Device Meeting (IEDM)*, S. 325 (1993)
- 3.9 T. Kuroi, S. Kusunoki, M. Shirahata, Y. Okumura, M. Kobayashi, M. Inuishi, N. Tsubouchi, The effects of nitrogen implantation into p⁺-polysilicon gate on gate oxide properties, *Proceedings of the 1994 Symposium on VLSI Technology*, S. 107 (1994)

- 3.10 S. C. Sun, C. H. Chen, D. L. W. Yen, C. J. Lin, Characterization and Optimization of NO-Nitrided Gate Oxide by RTP, *Technical Digest of the 1995 International Electron Device Meeting (IEDM)*, S. 687 (1995)
- 3.11 B. Maiti, P. J. Tobin, V. Misra, R. I. Hedge, K. G. Reid, C. Gelatos, High Performance 20Å NO Oxynitride for Gate Dielectric in Deep Sub-Quarter Micron CMOS Technology, *Technical Digest of the 1997 International Electron Device Meeting (IEDM)*, S. 651 (1997)
- 3.12 L. K. Han, M. Bhat, D. Wristers, H. H. Wang, D. L. Kwong, Recent Developments in Ultra Thin Oxynitride Gate Dielectrics, *Microelectronic Engineering* 28, S. 89 (1995)
- 3.13 T. Hori, H. Iwasaki, Improved transconductance under high normal field in MOSFET's with ultrathin nitrided oxides, *IEEE Electron Device Letters* 10, S. 195 (1989)
- 3.14 T. Hori, Inversion Layer Mobility under High Normal Field in Nitrided-Oxide MOSFET's, *IEEE Trans. on Electron Devices* 37(9), S. 2058 (1990)
- 3.15 C. Schlünder, R. Brederlow, P. Wiczorek, Claus Dahl, Jürgen Holz, Michael Röhner, Sylvia Kessel, Volker Herold, Karl Goser, Werner Weber, Roland Thewes, Trapping Mechanisms in Negative Bias Temperature Stressed p-MOSFETs, *Proceedings of the 1999 European Symposium on Reliability of Electron Devices, Failure Physics and Analysis (ESREF)* (1999)
- 3.16 T. Yamamoto, K. Uwasawa, T. Mogami, Bias Temperature Instability in Scaled p⁺ Polysilicon Gate p-MOSFET's, *IEEE Trans. on Electron Devices* 46(5), S. 921 (1999)
- 3.17 K. Uwasawa, T. Mogami, T. Kunio, M. Fukum, Scaling limitations of gate oxide in p⁺-polysilicon gate MOS structures for sub-quarter micron CMOS devices, *Technical Digest of the 1993 International Electron Device Meeting (IEDM)*, S. 895 (1993)

Kapitel 4

- 4.1 R. H. Fowler, L. Nordheim, Electron Emission in Intense Electric Fields, *Proc. of the Royal Society London Ser. A* 119, S. 173 (1928)
- 4.2 T. Hori, *Gate Dielectrics and MOS ULSIs*, Verlag: Springer, ISBN 3-540-63182-8 (1997)
- 4.3 J. G. Simmons, Generalized Formula for the Electric Tunnel Effect between Similar Electrodes Separated by a Thin Insulating Film, *Journal of Applied Physics* 34, S. 1793 (1963)
- 4.4 F. Schwabl, *Quantenmechanik (QM I)*, Verlag: Springer, ISBN 3-540-63779-6 (1998)
- 4.5 Y. Ando, T. Itoh, Calculation of transmission tunneling current across arbitrary potential, *Journal of Applied Physics* 61(4), S. 1497 (1987)
- 4.6 Persönliche Mitteilung von Prof. Dr. E. Vass an Dr. M. Kerber (Sept. 1999).
- 4.7 Werkstudentenbericht von G. Hechenblaikner über die Implementierung des Algorithmus von Ando et al. in ein Computerprogramm bei der Siemens AG, Bereich Halbleiter (Juli 1998).
- 4.8 G. P. Petersson, C. M. Svensson, J. Maserjian, Resonance effects observed at the onset of Fowler-Nordheim tunneling in thin MOS structures, *Solid-State Electronics* 18, S. 449 (1974)
- 4.9 A. Schenk, G. Heiser, Modeling and simulation of tunneling through ultra-thin gate dielectrics, *Journal of Applied Physics* 81(12), S. 7900 (1997)
- 4.10 B. De Salvo, R. Clerc, P. Masson, Y. P. Ahouassa, G. Ghibaudo, Electrical Characterization and Modeling of Ultra-thin (1.8-3.4 nm) Gate Oxides, *Proceedings of the 1999 European Solid State Device Research Conference (ESSDERC)*, S. 168 (1999)
- 4.11 S. M. Sze, *Physics of Semiconductor Devices*, Verlag: John Wiley & Sons, ISBN 0-471-05661-8 (1981)
- 4.12 R. F. Pierret, *Semiconductor Device Fundamentals*, Verlag: Addison-Wesley Publishing Company, ISBN 0-201-54393-1 (1996)
- 4.13 T. Pompl, M. Kerber, H. Wurzer, I. Eisele, Contribution of interface traps to valence band electron tunneling in PMOS devices, *Proceedings of the 2000 European Solid State Device Research Conference (ESSDERC)*, S. 292 (2000)

- 4.14 K. F. Schuegraf, C. Hu, Reliability of thin SiO₂, *Semiconductor Science and Technology* 9, S. 989 (1994)
- 4.15 E. Rosenbaum, L. F. Register, Mechanism of Stress-Induced Leakage Current in MOS Capacitors, *IEEE Trans. on Electron Devices* 44(2), S. 317 (1997)
- 4.16 D. J. DiMaria, E. Cartier, Mechanism for stress-induced leakage currents in thin silicon dioxide films, *Journal of Applied Physics* 78(6), S. 3883 (1995)
- 4.17 B. Ricco, M. Y. Azbel, M. H. Brodsky, Novel Mechanism for Tunneling and Breakdown of Thin SiO₂ Films, *Physical Review Letters* 51(19), S. 1795 (1983)
- 4.18 T. Nguyen, P. Olivo, B. Ricco, A new failure mode of very thin (<50 Å) thermal SiO₂ films, *Proceedings of the 1987 International Reliability Physics Symposium (IRPS)*, S. 66 (1987)
- 4.19 B. Ricco, G. Gozzi, M. Lanzoni, Modeling and Simulation of Stress-Induced Leakage Current in Ultrathin SiO₂ Films, *IEEE Trans. on Electron Devices* 45(7), S. 1554 (1998)
- 4.20 M. Kimura, H. Koyama, Stress-induced low level leakage mechanism in ultrathin silicon dioxide films caused by neutral oxide trap generation, *Proceedings of the 1994 International Reliability Physics Symposium (IRPS)*, S. 167 (1994)
- 4.21 M. Kimura, T. Ohmi, Conduction mechanism and origin of stress-induced leakage current in thin silicon dioxide films, *Journal of Applied Physics* 80(11), S. 6360 (1996).
- 4.22 S. Takagi, N. Yasuda, A. Toriumi, Experimental Evidence of Inelastic Tunneling and New I-V Model for Stress-induced Leakage Current, *Technical Digest of the 1996 International Electron Device Meeting (IEDM)*, S. 323 (1996)
- 4.23 D. J. Dumin, J. R. Maddux, Correlation of Stress-Induced Leakage Current in Thin Oxides with Trap Generation Inside the Oxides, *IEEE Trans. on Electron Devices* 40(5), S. 986 (1993)
- 4.24 C. Chang, C. Hu, R. Brodersen, Quantum yield of electron impact ionization in silicon, *Journal of Applied Physics* 57(2), S. 302 (1985)
- 4.25 B. Eitan, A. Kolodny, Two components of tunneling currents in metal-oxide-silicon structures, *Applied Physics Letters* 43(1), S. 106 (1983)
- 4.26 Z. A. Weinberg, M. V. Fischetti, Investigation of the SiO₂-induced substrate current in silicon field-effect transistors, *Journal of Applied Physics* 57(2), S. 443 (1985)
- 4.27 D. J. DiMaria, T. N. Theis, J. R. Kirtley, F. L. Pesavento, D. W. Dong, S. D. Brorson, Electron heating in silicon dioxide and off-stoichiometric silicon dioxide films, *Journal of Applied Physics* 57(4), S. 1214 (1985)
- 4.28 K. F. Schuegraf, C. Hu, Hole injection oxide breakdown model for very low voltage lifetime extrapolation", *Proceedings of the 1993 International Reliability Physics Symposium (IRPS)*, S. 7 (1993)

Kapitel 5

- 5.1 M. Depas, T. Nigam, M. M. Heyns, Soft Breakdown of Ultra-Thin Gate Oxide Layers, *IEEE Trans. on Electron Devices* 43(9), S. 1499 (1996)
- 5.2 S.-H. Lee, B.-J. Cho, J.-C. Kim, S.-H. Choi, Quasi-breakdown of ultrathin gate oxide under high field stress, *Technical Digest of the 1994 International Electron Device Meeting (IEDM)*, S. 605 (1994)
- 5.3 A. Halimaoui, O. Brière, G. Ghibaudo, Quasi-breakdown in ultrathin gate dielectrics, *Microelectronic Engineering* 36, S. 157 (1997)
- 5.4 R. Moazzami, C. Hu, Stress-induced current in thin silicon dioxide films, *Technical Digest of the 1992 International Electron Device Meeting (IEDM)*, S. 139 (1992)
- 5.5 E. F. Runnion, S. M. Gladstone, R. S. Scott, D. J. Dumin, L. Lie, J. C. Mitros, Thickness Dependence of Stress-Induced Leakage Currents in Silicon Oxide, *IEEE Trans. on Electron Devices* 44(6), S. 993 (1997)

- 5.6 K. Naruke, S. Taguchi, M. Wada, Stress-induced leakage current limiting to scale down EEPROM tunnel oxide thickness, *Technical Digest of the 1988 International Electron Device Meeting (IEDM)*, S. 424 (1988)
- 5.7 E. Rosenbaum, L. F. Register, Mechanism of Stress-Induced Leakage Current in MOS Capacitors, *IEEE Trans. on Electron Devices* 44(2), S. 317 (1997)
- 5.8 N. K. Patel, A. Toriumi, Stress-induced leakage current in ultrathin SiO₂ films, *Applied Physics Letters* 64(14), S. 1809 (1994)
- 5.9 D. J. DiMaria, E. Cartier, Mechanism for stress-induced leakage currents in thin silicon dioxide films, *Journal of Applied Physics* 78(6), S. 3883 (1995)
- 5.10 D. J. DiMaria, D. Arnold, E. Cartier, Degradation and breakdown of silicon dioxide films on silicon, *Applied Physics Letters* 61(19), S. 2329 (1992)
- 5.11 O. Brière, A. Halimaoui, G. Ghibaudo, Breakdown characteristics of ultra thin gate oxides following field and temperature stresses, *Solid-State Electronics* 41(7), S. 981 (1997)
- 5.12 T. Pompl, H. Wurzer, M. Kerber, R.C.W. Wilkins, I. Eisele, Influence of Soft Breakdown on NMOSFET Device Characteristics, *Proceedings of the 1999 International Reliability Physics Symposium (IRPS)*, S. 82 (1999)
- 5.13 F. Crupi, R. Degraeve, G. Groeseneken, T. Nigam, H. E. Maes, On the Properties of the Gate and Substrate Current after Soft Breakdown in Ultrathin Oxide Layers, *IEEE Trans. on Electron Devices* 45(11), S. 2329 (1998)
- 5.14 K. Okada, S. Kawasaki, Y. Hirofujii, New experimental findings on stress induced leakage current of ultra thin silicon dioxides, *Extended Abstracts of Solid State Devices and Materials*, S. 565 (1994)
- 5.15 O. Brière, J. A. Chroboczek, G. Ghibaudo, Random telegraph signal in the quasi-breakdown current of MOS capacitors, *Proceedings of the 1996 European Solid State Device Research Conference (ESSDERC)*, S. 759 (1996)
- 5.16 E. Miranda, J. Suñé, R. Rodriguez, M. Nafria, X. Aymerich, Switching Behavior of the Soft Breakdown Conduction Characteristic in Ultra-thin (<5 nm) Oxide MOS Capacitors, *Proceedings of the 1998 International Reliability Physics Symposium (IRPS)*, S. 42 (1998)
- 5.17 M. Houssa, N. Vandewalle, T. Nigam, M. Ausloos, P. W. Mertens, M. M. Heyns, Analysis of the Gate Voltage Fluctuations in Ultra-Thin Gate Oxides After Soft Breakdown, *Technical Digest of the 1998 International Electron Device Meeting (IEDM)* (1998)
- 5.18 B. E. Weir, P. J. Silverman, D. Monroe, K. S. Kirsch, M. A. Alam, G. B. Alers, T. W. Sorsch, G. L. Timp, F. Baumann, C. T. Liu, Y. Ma, D. Hwang, Ultra-Thin Gate Dielectrics: They Break Down, But Do They Fail?, *Technical Digest of the 1997 International Electron Device Meeting (IEDM)*, S. 73 (1997)
- 5.19 T. Pompl, H. Wurzer, M. Kerber, I. Eisele, Influence of gate oxide breakdown on MOSFET device operation, *Microelectronics Reliability* 40(1), S. 37-47 (2000)
- 5.20 E. Harari, Dielectric breakdown in electrically stressed thin films of thermal SiO₂, *Journal of Applied Physics* 49(4), S. 2478 (1978)
- 5.21 D. J. Dumin, Wearout and Breakdown in Thin Silicon Oxide, *Journal of the Electrochemical Society* 142(4), S. 1272 (1995)
- 5.22 R. Degraeve, G. Groeseneken, R. Bellens, M. Depas, H. E. Maes, A consistent model for the thickness dependence of intrinsic breakdown in ultra-thin oxides, *Technical Digest of the 1995 International Electron Device Meeting (IEDM)*, S. 863 (1995)
- 5.23 R. Degraeve, G. Groeseneken, R. Bellens, J. L. Ogier, M. Depas, P. J. Roussel, H. E. Maes, New Insights in the Relation Between Electron Trap Generation and the Statistical Properties of Oxide Breakdown, *IEEE Trans. on Electron Devices* 45(4), S. 904 (1998)
- 5.24 J. Stathis, D. J. DiMaria, Reliability Projection for Ultra-Thin Oxides at Low Voltage, *Technical Digest of the 1998 International Electron Device Meeting (IEDM)* (1998)
- 5.25 T. Nigam, R. Degraeve, G. Groeseneken, M. M. Heyns, H. E. Maes, Constant Current Charge-to-breakdown: still a valid tool to study the reliability of MOS structures?, *Proceedings of the 1998 International Reliability Physics Symposium (IRPS)*, S. 62 (1998)

- 5.26 D. J. DiMaria, T. N. Theis, J. R. Kirtley, F. L. Pesavento, D. W. Dong, S. D. Brorson, Electron heating in silicon dioxide and off-stoichiometric silicon dioxide films, *Journal of Applied Physics* 57(4), S. 1214 (1985)
- 5.27 M. V. Fischetti, D. J. DiMaria, L. Dori, J. Batey, E. Tierney, J. Stasiak, Ballistic electron transport in thin silicon dioxide films, *Physical Review B* 35(9), S. 4404 (1987)
- 5.28 A. v. Schwerin, W. Bergner, H. Jacobs, Self-Consistent Simulation of Hot-Carrier Damage Enhanced Gate Induced Drain Leakage, *Technical Digest of the 1992 International Electron Device Meeting (IEDM)*, S. 543 (1992)
- 5.29 J. Maserjian, N. Zamani, Observation of positively charged state generation near the Si/SiO₂ interface during Fowler-Nordheim tunneling, *Journal of Vacuum Science and Technology* 20(3), S. 743 (1982)
- 5.30 T. Hori, *Gate Dielectrics and MOS ULSIs*, Verlag: Springer, ISBN 3-540-63182-8 (1997)
- 5.31 R. Degraeve, N. Pangon, B. Kaczer, T. Nigam, G. Groeseneken, Temperature acceleration of oxide breakdown and its impact on ultra-thin gate oxide reliability, *Proceedings of the 1999 Symposium on VLSI Technology* (1999)
- 5.32 M. Depas, T. Nigam, M. M. Heyns, Definition of dielectric breakdown for ultra thin (<2 nm) gate oxides, *Solid-State Electronics* 41(5), S. 7 (1997)
- 5.33 E. Wu, E. Nowak, J. Aitken, W. Abadeer, L. K. Han, S. Lo, Structural Dependence of Dielectric Breakdown in Ultra-Thin Gate Oxides and Its Relationship to Soft Breakdown Modes and Device Failure, *Technical Digest of the 1998 International Electron Device Meeting (IEDM)* (1998)
- 5.34 T. Pompl, H. Wurzer, M. Kerber, I. Eisele, Investigation of ultra-thin gate oxide reliability behavior by separate characterization of soft breakdown and hard breakdown, *Proceedings of the 2000 International Reliability Physics Symposium (IRPS)*, S. 40 (2000)
- 5.35 J. McPherson, H. Mogul, Disturbed bonding states in SiO₂ thin-films and their impact on time-dependent dielectric breakdown, *Proceedings of the 1998 International Reliability Physics Symposium (IRPS)*, S. 47 (1998)
- 5.36 M. Kimura, Field and Temperature Acceleration Model for Time-Dependent Dielectric Breakdown, *IEEE Trans. on Electron Devices* 46(1), S. 220 (1999),
- 5.37 J. S. Suehle, P. Chaparala, C. Messick, W. M. Miller, K. C. Boyko, Field and Temperature Acceleration of Time-Dependent Dielectric Breakdown in Intrinsic Thin SiO₂, *Proceedings of the 1994 International Reliability Physics Symposium (IRPS)*, S. 120 (1994)
- 5.38 S. Suehle, P. Chaparala, Low Electric Field Breakdown of Thin SiO₂ Films Under Static and Dynamic Stress, *IEEE Trans. on Electron Devices* 44(5), S. 801 (1997)
- 5.39 A. Yassine, H. E. Nariman, K. Olasupo, "Field and Temperature Dependence of TDDb of Ultrathin Gate Oxide", *IEEE Electron Device Letters* 20(8), p. 390 (1999)
- 5.40 S. Bruyère, D. Roy, E. Vincent, G. Ghibaudo, Temperature Acceleration of Breakdown and Quasi-Breakdown Phenomena in Ultra-thin Oxides, *Proceedings of the 1999 European Symposium on Reliability of Electron Devices, Failure Physics and Analysis (ESREF)* (1999)
- 5.41 D. J. DiMaria, J. H. Stathis, Non-Arrhenius temperature dependence of reliability in ultrathin silicon dioxide films, *Applied Physics Letters* 74(12), S. 1752 (1999)
- 5.42 B. Kaczer, R. Degraeve, N. Pangon, T. Nigam, G. Groeseneken, The effect of elevated temperature on the reliability of very thin oxide films, *Proceedings of the 1999 European Solid State Device Research Conference (ESSDERC)*, S. 356 (1999)
- 5.43 S. M. Sze, *VLSI Technology*, Second Edition, Chapt. 7.9, Tab. 7, S. 315, Verlag: McGraw-Hill Book Company, ISBN 0-07-100347-9 (1988)
- 5.44 D. L. Griscom, Diffusion of radiolytic molecular hydrogen as a mechanism for the post-irradiation buildup of interface states in SiO₂-on-Si structures, *Journal of Applied Physics* 58(7), S. 2524 (1985)
- 5.45 D. R. Lide, *CRC Handbook of Chemistry and Physics*, 73rd Edition, Chap. 9-131, Verlag: CRC Press Inc. (1993)
- 5.46 H. S. Momose, M. Ono, T. Yoshitomi, T. Ohguro, S. Nakamura, M. Saito, H. Iwai, 1.5 nm Direct-Tunneling Gate Oxide Si MOSFET's, *IEEE Trans. on Electron Devices* 43(8), S. 1233 (1996)

Kapitel 6

- 6.1 T. P. Ma, Making Silicon Nitride Film a Viable Gate Dielectric, *IEEE Trans. on Electron Devices* 45(3), S. 680 (1998)
- 6.2 K. Sekine, Y. Saito, M. Hirayama, T. Ohmi, High-integrity Ultra-thin Silicon Nitride Film Grown at Low Temperature for Extending Scaling Limit of Gate Dielectric, *Proceedings of the 1999 Symposium on VLSI Technology* (1999)
- 6.3 S. Mahapatra, V. R. Rao, K. N. ManjulaRani, C. D. Parikh, J. Vasi, B. Cheng, M. Khare, J. C. S. Woo, 100 nm Channel Length MNSFETs using a Jet Vapor Deposited Ultra-thin Silicon Nitride Gate Dielectric, *Proceedings of the 1999 Symposium on VLSI Technology* (1999)
- 6.4 S. M. Sze, *Physics of Semiconductor Devices*, Verlag: John Wiley & Sons, ISBN 0-471-05661-8 (1981)
- 6.5 I. C. Kizilyalli, R. Y. S. Huang, MOS Transistors with Stacked $\text{SiO}_2\text{-Ta}_2\text{O}_5\text{-SiO}_2$ Gate Dielectric for Giga-Scale Integration of CMOS Technologies, *IEEE Electron Device Letters* 19(11), S. 423 (1998)
- 6.6 D. Park, Q. Lu, T. King, C. Hu, A. Kalnitsky, S. Tay, C. Cheng, $\text{SiON/Ta}_2\text{O}_5\text{/TiN}$ Gate-Stack Transistor with 1.8nm Equivalent SiO_2 Thickness, *Technical Digest of the 1998 International Electron Device Meeting (IEDM)* (1998)
- 6.7 T. Devoivre, C. Papadas, M. Setton, A Reliable 0.1 μm Ta_2O_5 Transistor Manufactured with an Almost Standard CMOS Process, *Proceedings of the 1999 Symposium on VLSI Technology* (1999)
- 6.8 Y. Momiyama, H. Minakata, T. Sugii, Ultra-Thin $\text{Ta}_2\text{O}_5\text{/SiO}_2$ Gate Insulator with TiN Gate Technology for 0.1 μm MOSFETs, *Proceedings of the 1997 Symposium on VLSI Technology*, S. 135 (1997)
- 6.9 A. Chin, C. C. Liao, C. H. Lu, W. J. Chen, C. Tsai, Device and Reliability of High-K Al_2O_3 Gate Dielectric with Good Mobility and Low D_{it} , *Proceedings of the 1999 Symposium on VLSI Technology* (1999)
- 6.10 L. Manchanda, W. H. Lee, J. E. Bower, F. H. Baumann, W. L. Brown, C. J. Case, R. C. Keller, Y. O. Kim, E. J. Laskowski, M. D. Morris, R. L. Opila, P. J. Silverman, T. W. Sorsch and G. R. Weber, Gate Quality Doped High K Films for CMOS Beyond 100 nm: 3 - 10nm Al_2O_3 with Low Leakage and Low Interface States, *Technical Digest of the 1998 International Electron Device Meeting (IEDM)* (1998)
- 6.11 S. A. Campbell, D. C. Gilmer, X. Wang, M. Hsieh, H. Kim, W. L. Gladfelter, J. Yan, MOSFET Transistors Fabricated with High Permittivity TiO_2 Dielectrics, *IEEE Trans. on Electron Devices* 44(1), S. 104 (1997)
- 6.12 X. Guo, T. P. Ma, T. Tamagawa, B. L. Halpern, High Quality Ultra-thin $\text{TiO}_2\text{/Si}_3\text{N}_4$ Gate Dielectric for Giga Scale MOS Technology, *Technical Digest of the 1998 International Electron Device Meeting (IEDM)* (1998)
- 6.13 C. Hobbs, R. Hegde, B. Maiti, H. Tseng, D. Gilmer, P. Tobin, O. Adetutu, F. Huang, D. Weddington, R. Nagabushnam, D. O'Meara, K. Reid, L. La, L. Grove, M. Rossow, Sub-Quarter Micron CMOS Process for TiN-Gate MOSFETs with TiO_2 Gate Dielectric formed by Titanium Oxidation, *Proceedings of the 1999 Symposium on VLSI Technology* (1999)
- 6.14 H. Landolt, R. Börnstein, K.-H. Hellwege, O. Madelung, *Zahlenwerte und Funktionen aus Naturwissenschaften und Technik, Gruppe 3: Kristall- und Festkörperphysik Bd. 17 Halbleiter Teilband g: Physik der nicht-tetraedrisch gebundenen binären Verbindungen 3*, Kapitel 9.15.2.1.1, Verlag: Springer, ISBN 3-540-12744-5 (1984)
- 6.15 H. Treichel, A. Mitwalsky, G. Tempel, G. Zorn, D. A. Bohling, K. R. Coyle, B. S. Felker, M. George, W. Kern, A. P. Lane, N. P. Sandler, Deposition, Annealing and Characterisation of High-dielectric-constant Metal Oxide Films, *Advanced Materials for Optics and Electronics* 5, S. 163 (1995)
- 6.16 H. Kim, D. C. Gilmer, S. A. Campbell, D. L. Polla, Leakage current and electrical breakdown in metal-organic chemical vapor deposited TiO_2 dielectrics on silicon substrates, *Applied Physics Letters* 69(25), S. 3860 (1996)

- 6.17 S. C. Sun, T. F. Chen, A Novel Approach for Leakage Current Reduction of LPCVD Ta_2O_5 and TiO_2 Films by Rapid Thermal N_2O Annealing, *Technical Digest of the 1994 International Electron Device Meeting (IEDM)*, S. 333 (1994)
- 6.18 Deutsche Patentanmeldung 19856084.2, *Verfahren zur Herstellung einer Metalloxidschicht bzw. einer strukturierten Metalloxidschicht* (1998)
- 6.19 M. M. Moslehi, K. C. Saraswat, Thermal Nitridation of Si and SiO_2 for VLSI, *IEEE Trans. on Electron Devices ED-32(2)*, S. 106 (1985)
- 6.20 Deutsche Patentanmeldung 19856082.6, *Verfahren zur Strukturierung einer Metalloxidschicht* (1999)
- 6.21 S. M. Sze, *VLSI Technology*, Verlag: McGraw-Hill, ISBN 0-07-100347-9 (1988)
- 6.22 S. Pindl, *Entwicklung von CMOS-Transistoren auf Silicon-on-Insulator-Substraten*, Dissertation (1998)

Anhang A

- 7.1 T. Hori, *Gate Dielectrics and MOS ULSIs*, Verlag: Springer, ISBN 3-540-63182-8 (1997)
- 7.2 S. Takagi, A. Toriumi, M. Iwase, H. Tango, On the Universality of Inversion Layer Mobility in Si MOSFET's: Part II – Effects of Surface Orientation, *IEEE Trans. on Electron Devices 41(12)*, S. 2363 (1994)
- 7.3 G. J. Hu, Design Tradeoffs Between Surface and Buried-Channel FET's, *IEEE Trans. on Electron Devices ED-32(3)*, S. 584 (1985)
- 7.4 S. M. Sze, *Physics of Semiconductor Devices*, Verlag: John Wiley & Sons, ISBN 0-471-05661-8 (1981)

Anhang B

- 8.1 R. Storm, *Wahrscheinlichkeitsrechnung, mathematische Statistik und statistische Qualitätskontrolle*, Verlag: Fachbuchverlag Leipzig, ISBN 3-343-00871-0 (1995)
- 8.2 E. Gumbel, *Statistics of Extremes*, 4th edition, Verlag: Columbia University Press (1967)

Verzeichnis der Abkürzungen

Abkürzung	Beschreibung
BTS	„Bias Temperature Stress“
CHE	„Channel Hot-Electron Injection“
CMOS	„Complementary Metal Oxide Semiconductor“
CP	„Charge Pumping“
DAHC	„Drain Avalanche Hot-Carrier Injection“
DIBL	„Drain-Induced Barrier Lowering“
DRAM	„Dynamic Random-Access Memory“
DT	Direktes Tunneln
FET	Feldeffekttransistor („Field Effect Transistor“)
FIT	Allgemein gültige Fehlereinheit („failure unit“)
FN	Fowler-Nordheim
GI	Gateinjektion („gate injection“)
GIDL	„Gate-Induced Drain Leakage“
G.O.	Größenordnung
HBD	„Hard Breakdown“
HCS	„Hot Carrier Stress“
HF-CV	Hochfrequenz-CV („high frequency CV“)
HRD	„Hydrogen Related Defect“
JVD	„Jet Vapor Deposition“
LDD	„Lightly Doped Drain“
LF-CV	Niederfrequenz-CV („low frequency CV“)
LPCVD	„Low Pressure Chemical Vapor Deposition“
MOCVD	„Metal Organic Chemical Vapor Deposition“
MOS	„Metal Oxide Semiconductor“
MOSFET	„Metal Oxide Semiconductor Field Effect Transistor“
n	n-dotiertes Gebiet
n ⁺	Entartet n-dotiertes Gebiet
NBTS	„Negative Bias Temperature Stress“
NMOSFET	MOSFET vom Typ n-Kanal
OED	„Oxidation-Enhanced Diffusion“

ORD	„Oxidation-Retarded Diffusion“
p	p-dotiertes Gebiet
p ⁺	Entartet p-dotiertes Gebiet
PECVD	„Plasma Enhanced Chemical Vapor Deposition“
PLD	„Peroxy Linkage Defect“
PMOSFET	MOSFET vom Typ p-Kanal
ppm	„parts per million“
RTA	„Rapid Thermal Anneal“
RTP	„Rapid Thermal Process“
SBD	„Soft Breakdown“
SGHE	„Secondarily Generated Hot-Electron Injection“
SHE	„Substrate Hot-Electron Injection“
SHI	„Substrate Hot-Hole Injection“
SI	Substratinjektion („substrate injection“)
SIMS	Sekundärionenmassenspektrometrie („Secondary-Ion Mass Spectrometry“)
SILC	„Stress-Induced Leakage Current“
SMU	„Source Measurement Unit“
SRAM	„Static Random-Access Memory“
STI	„Shallow Trench Isolation“
TAT	„Trap-Assisted Tunneling“
TDDB	„Time Dependent Dielectric Breakdown“
TEM	Transmissionselektronenmikroskop
WKB	Wentzel-Kramers-Brillouin

Verzeichnis der Symbole

Symbol	Beschreibung	Einheit
\hbar	$\hbar/(2\pi)$	Js
α	„duty cycle“ eines Spannungspulses	
β	Formparameter in der Weibull-Verteilung („Weibull slope“)	
$\Delta\Phi_b$	Differenz der Barrierenhöhen von Polysilizium und Siliziumsubstrat gegenüber SiO_2	V
ΔH_0	Enthalpie für einen dielektrischen Durchbruch in SiO_2	J
Δx_{ox}	Oxiddünnung durch einen Defekt	nm
ϵ_0	Dielektrizitätskonstante im Vakuum: $8.854 \cdot 10^{-12}$	F/m
ϵ_{DH}	Dielektrizitätskonstante eines Dielektrikums	F/m
ϵ_{ox}	Dielektrizitätskonstante von Siliziumdioxid: $\epsilon_{ox} = 3.9 \cdot \epsilon_0$	F/m
ϵ_s	Dielektrizitätskonstante eines Halbleiters	F/m
ϵ_{Si}	Dielektrizitätskonstante von Silizium: $\epsilon_{Si} = 11.9 \cdot \epsilon_0$	F/m
Φ_b	Potentialbarriere zwischen den Leitungsbandkanten von Halbleiter und Isolator	V
$\Phi_{b,m}$	Potentialbarriere zwischen den Leitungsbandkanten von Metallelektrode und Isolator	V
Φ_{ms}	Austrittsarbeitsdifferenz zwischen Gate und Siliziumsubstrat	V
Γ	Gammafunktion	
γ	Feldbeschleunigungsfaktor	cm/MV
φ_i	Minimale Elektronenenergie zur Erzeugung einer Stoßionisation	eV
φ_{it}	Kritische Elektronenenergie für die Erzeugung eines Grenzflächenzustandes	eV
λ	Fehlerrate	
μ	Erwartungswert	
μ_{eff}	Effektive Beweglichkeit	cm^2/Vs
μ_{FE}	Feldeffektbeweglichkeit	cm^2/Vs
π	Kreiszahl: $\approx 3,1416$	
ρ	Ladungsträgerdichte	C/cm^2
σ	Streuung	
σ_n	Einfangsquerschnitt für Elektronen	cm^2
σ_p	Einfangsquerschnitt für Löcher	cm^2

σ_T	Wirkungsquerschnitt einer Haftstelle („trap“) in SiO ₂	cm ²
τ	Lebensdauer	s
τ_{it}	Zeitkonstante für die Besetzung von Grenzflächenzuständen	s
ω	Kreisfrequenz	s ⁻¹
Ψ_B	Substratpotential	V
Ψ_j	Wellenfunktion in der j-ten Potentialstufe	
Ψ_{min}	Minimales Potential im „Buried Channel“-Gebiet	V
Ψ_s	Bandverbiegung an der Oberfläche des Siliziumsubstrates	V
A	Fläche	cm ²
B	Koeffizient im Exponenten der Fowler-Nordheim-Tunnelstromdichte	MV/cm
C	Kapazität	F
C _{bc}	Kapazität der Raumladungszone des „Buried Channel“-Gebietes	F
C _D	Kapazität der Verarmungszone	F
C _{DIF}	Gesamte Flächenkapazität eines Dielektrikums	F/m ²
C _{HF}	Kapazität eines MOS-Systems bei einer CV-Messung mit hohen Frequenzen	F
C _I	Kapazität der Inversionsladungen	F
C _{it}	Kapazität der Grenzflächenzustände	F
C _I	Flächenkapazität einer einzelnen Schicht eines Dielektrikums aus mehreren Schichtfolgen	F/m ²
Cl ⁻	Einfach ionisiertes Chlor	
C _{LF}	Kapazität eines MOS-Systems bei einer CV-Messung mit niedrigen Frequenzen	F
C _{ox}	Gateoxiddkapazität	F
C _s	Kapazität im Silizium	F
D	Defektdichte	cm ⁻²
D _{it}	Energieabhängige Grenzflächenzustandsdichte	eV ⁻¹ cm ⁻²
E	Energie	J
e	Exponentialzahl: $\approx 2,71828$	
e ⁻	Symbol für ein Elektron	
E _a	Aktivierungsenergie	eV
E _c	Energieniveau der Leitungsbandkante eines Halbleiters	eV
E _{eff}	Effektives elektrisches Feld	MV/cm
E _F	Fermi-niveau	eV
E _g	Elektrisches Feld zwischen Gate und Gegenelektrode	MV/cm
E _{g,ox}	Bandlücke eines Oxides	eV
E _{gap}	Bandlücke eines Halbleiters	eV
E _i	Intrinsisches Energieniveau eines Halbleiters	eV
E _{ox}	Elektrisches Feld im Oxid	MV/cm
E _{pot}	Potentielle Energie, z.B. von Elektronen	eV
E _v	Energieniveau der Valenzbandkante eines Halbleiters	eV
F	Verteilungsfunktion (auch Ausfallanteil oder Fehleranteil genannt)	

f	Frequenz	Hz
f_{duty}	„duty factor“	
G	Koeffizient im I/E-Modell: $G = B + H$	MV/cm
g_d	Kanalleitfähigkeit	S
g_m	Transferleitfähigkeit	S
h	Plancksches Wirkungsquantum: $6.626 \cdot 10^{-34}$	Js
H	Koeffizient im Exponenten für die Generation von Löchern	MV/cm
H^+	Ionisiertes Wasserstoffatom	
h^+	Symbol für ein Loch	
HCl	Salzsäure	
i	Imaginäre Einheit	
I_{CP}	Charge-Pumping-Strom	A
I_d	Drainstrom	A
$I_{d,\text{lin}}$	Drainstrom bei $V_g = V_{D0}$ im linearen Bereich des Transistors	A
$I_{d,\text{sat}}$	Drainstrom bei $V_g = V_{D0}$ im Sättigungsbereich des Transistors	A
I_{Drain}	Drainstrom	A
I_g	Gateleckstrom	A
I_{Gate}	Gateleckstrom	A
I_{GIDL}	„Gate-Induced Drain Leakage“-Strom	A
I_{off}	Off-Strom eines Transistors	A
I_s	Sourcestrom	A
I_{Source}	Sourcestrom	A
I_{sub}	Substratstrom	A
I_t	Drainstrom bei der Einsatzspannung eines Transistors	A
I_{Wanne}	Strom am Wannenkontakt	A
J	Stromdichte	A/cm ²
J_{Diff}	Am Diffusionskontakt gemessene Stromdichte	A/cm ²
J_{FN}	Fowler-Nordheim-Tunnelstromdichte	A/cm ²
J_g	Am Gate gemessene Stromdichte	A/cm ²
J_{Gate}	Am Gate gemessene Stromdichte	A/cm ²
$J_{\text{Gate,T}}$	Stromdichte normal tunnelnder Elektronen, d.h. ohne der Unterstützung von Haftstellen im Oxid	A/cm ²
$J_{\text{Gate,TAT}}$	Stromdichte an Elektronen, die mit der Unterstützung von Haftstellen im Oxid tunneln („Trap-Assisted Tunneling“)	A/cm ²
J_{inj}	Stromdichte, die in ein Oxid injiziert wurde	A/cm ²
J_{Wanne}	Am Wannenkontakt gemessene Stromdichte	A/cm ²
k	Boltzmannkonstante: $1,381 \cdot 10^{-23}$	J/K
k_j	Betrag des k-Vektors in der j-ten Potentialstufe	
L	Länge eines Transistors	μm
L_{ov}	Laterale Überlapplänge von Gate und Diffusionsgebiet	μm
m^*	Effektive Masse	kg

m_e	Masse eines freien Elektrons: $9,11 \cdot 10^{-31}$	kg
m_j	Effektive Masse eines Elektrons in der j-ten Potentialstufe	kg
m_{ox}	Effektive Masse eines Elektrons in der Bandlücke von SiO_2	kg
N_A	Konzentration an Akzeptoren	cm^{-3}
N_D	Konzentration an Donatoren	cm^{-3}
N_{dop}	Dotierstoffkonzentration	cm^{-3}
n_i	Intrinsische Ladungsträgerkonzentration von Silizium: $1,5 \cdot 10^{10}$ (bei $27^\circ C$)	cm^{-3}
N_{it}	Dichte an Grenzflächenzuständen	cm^{-2}
N_{poly}	Dotierung im Polysilizium	cm^{-3}
N_{sub}	Substratdotierung	cm^{-3}
$O\cdot$	Gebundenes Sauerstoffatom mit einer freien Bindung	
$O\cdot\cdot$	Gebundenes Sauerstoffatom mit einer freien Bindung und einem zusätzlichen Elektron	
p_k	Ereigniswahrscheinlichkeit	
q	Elementarladung	C
Q_{bc}	Ladung im vollständig verarmten „Buried Channel“-Gebiet	C
Q_{BD}	Ladungsdichte bis zum Oxiddurchbruch	C/cm^2
Q_D	Ladung in der Verarmungszone	C
Q_f	Fixe Oxidladung	C
Q_i	Inversionsladung	C
Q_{it}	Grenzflächenladungsdichte	C/cm^2
Q_{ss}	Gepumpte Ladung pro Puls	C
R	Ohmscher Widerstand	Ω
r	Abstand zu einer Punktladung	m
S	„subthreshold swing“	mV/ Dekade
$Si\cdot$	Gebundenes Siliziumatom mit einer freien Bindung	
$Si\cdot\cdot$	Gebundenes Siliziumatom mit einer freien Bindung und einem zusätzlichen Elektron	
T	Periode	s
T	Temperatur	K
t_{active}	Gesamtdauer der elektrischen Belastung einer aktiven Gateoxidfläche	s
t_{BD}	Zeit bis zum Oxiddurchbruch	s
T_{DIE}	Dicke eines Dielektrikums	nm
$t_{em, e}$	Emissionszeit für Elektronen	s
$t_{em, h}$	Emissionszeit für Löcher	s
t_f	Abfallzeit eines Spannungspulses	s
t_{HBD}	Zeit bis zum dielektrischen Durchbruch („Hard Breakdown“)	s
t_{high}	Zeit eines Spannungspulses im oberen Niveau	s
t_{low}	Zeit eines Spannungspulses im Grundniveau	s
T_{ox}	Oxiddicke	nm

t_r	Anstiegszeit eines Spannungspulses	s
t_{SBD}	Zeit bis zum „Soft Breakdown“	s
t_{use}	Betriebsdauer eines elektrischen Bauelementes	s
U_j	Höhe der j-ten Potentialstufe	V
U_p	Potential einer Punktladung	V
V_{bottom}	Grundniveau eines Spannungspulses	V
V_d	Drainspannung	V
$V_{d,\text{sat}}$	Sättigungsdrainspannung	V
V_{DD}	Versorgungsspannung (Betriebsspannung einer Technologie)	V
V_{fb}	Flachbandspannung	V
V_g	Gatespannung	V
V_{ox}	Oxidspannung	V
V_{poly}	Bandverbiegung im Polysiliziumgate	V
V_{Si}	Bandverbiegung im Siliziumsubstrat	V
V_{sub}	Substratspannung	V
V_i	Einsatzspannung	V
$V_{i,\text{lin}}$	Einsatzspannung im linearen Bereich eines Transistors	V
$V_{i,\text{sat}}$	Einsatzspannung im Sättigungsbereich eines Transistors	V
v_{th}	Elektronendriftgeschwindigkeit	cm/s
V_{top}	Oberes Niveau eines Spannungspulses	V
W	Weite eines Transistors	μm
w	Weite einer Verarmungszone	μm
w_m	Maximale Breite einer Verarmungszone	μm
X^-	Symbol für eine negativ geladene Haftstelle in SiO_2	
X^+	Symbol für eine positiv geladene Haftstelle in SiO_2	
X^0	Symbol für eine neutrale Haftstelle in SiO_2	
x_{bc}	Tiefe des „Buried Channel“-Gebietes	μm
x_{eff}	Effektive Oxiddicke im I/E-Modell: $x_{\text{eff}} = x_{ox} - \Delta x_{ox}$	nm
x_{min}	Ort des minimalen Potentials im „Buried Channel“-Gebiet	μm
x_{ox}	Oxiddicke	nm
x_p	x-Koordinate einer Punktladung	
Y	„Yield“ (Ausbeute)	
Z	Impedanz	Ω
z_p	z-Koordinate einer Punktladung	

Danksagung

Herrn Prof. Dr. I. Eisele möchte ich für die Überlassung des Themas herzlich danken. Zudem danke ich Herrn Prof. Dr. I. Eisele, daß er mir ermöglichte, ein Promotionsstipendium beim Halbleiterbereich der Siemens AG, seit dem 01.04.1999 Infineon Technologies AG, wahrzunehmen. Das Interesse an der Arbeit und die gewährten Freiheiten in der Gestaltung waren sehr wichtig für den Erfolg.

Ein besonderer Dank gilt Herrn Dr. H. Wurzer, der mit immerwährendem Engagement und steter Offenheit zur Diskussion sehr am Gelingen der Arbeit beteiligt war. Die Organisation von prozessiertem Silizium durch Herrn Dr. H. Wurzer war die grundlegende Voraussetzung für die experimentellen Ergebnisse.

Einen weiteren besonderen Dank möchte ich an Herrn Dr. M. Kerber richten, der mir eine große Unterstützung auf dem Gebiet der Meßtechnik zu Teil werden ließ. Außerdem war Herr Dr. M. Kerber bei fachlichen Diskussionen stets eine sehr große Hilfe.

Herrn Dr. E. Hammerl und Herrn Dr. D. Schumann danke ich sehr, daß sie in der Anfangsphase der Arbeit bei Fachfragen, bei organisatorischen Angelegenheiten und in Diskussionen über persönliche Entscheidungen geholfen haben. Beiden gilt der Dank auch für ihre stete Bereitschaft, sich für die Belange der Doktoranden einzusetzen.

Frau A. Rodemann und Herrn Dr. A. v. Schwerin möchte ich für die sehr hilfreichen Simulationen von „Gate-Induced Drain Leakage“ und Einsatzspannungen danken.

Herrn D. Takacs danke ich für die sehr hilfreichen Diskussionen über „Gate-Induced Drain Leakage“ und die exakte Erörterung weiterer Fachfragen.

Herrn W. Zatsch möchte ich für jegliche Unterstützung und die sehr gute Zusammenarbeit in Sachen Meßtechnik und Meßstationen danken.

Allen Mitgliedern der ehemaligen CMOS-Gruppe von ZTME I danke ich für die Unterstützung bei der Prozessierung von Losen und die gute Zusammenarbeit. Ein besonderes Dankeschön gilt dabei Herrn Dr. S. Pindl für das sehr kollegiale „time sharing“ seines Computers und hilfreiche Diskussionen.

Herrn Dr. A. Mattheus und Frau G. Innertsberger danke ich für die Hilfe bei der RTO-Prozessierung in München Perlach.

Ein herzlicher Dank gilt auch der Ofentechnik in München Perlach und der Ofentechnik in Dresden für die sehr gute Zusammenarbeit und die großzügige Unterstützung. Explizit

möchte ich mich in München Perlach bei Herrn S. Degel, Herrn M. Kaufmann, Herrn W. Sturm und Herrn P. Schiller bedanken. In Dresden gilt der Dank besonders Herrn J. Schulze, Herrn W. Rebitzer und Herrn T. Gärtner.

Für die außergewöhnliche Unterstützung beim „Experiment“ TiO_2 möchte ich mich bei allen Beteiligten sehr herzlich bedanken: Herr Dr. H. Wurzer (Prozeßintegration), Herr J. Schulze, Herr W. Rebitzer und Herr T. Gärtner (jeweils Oxidation), Herr W. Kegel (TiO_x), Herr M. Förster (Poly), Herr Dr. L. Dittmar (Gate), Frau K. Krah und Herr S. Wege (jeweils TiO_x -Ätzung), Frau K. Schupke (Reinigung) und Herr S. Schmidbauer (Ti-Sputtern).

Frau Dr. K. Wirth danke ich für die Einweisung zur selbstständigen Bedienung der Photoemissionsmikroskopie.

Frau I. Noack, Frau B. Tippelt, Herrn R. Lehmann, Frau V. Klüppel, Herrn R. Lemme und Herrn Dr. F. Jahnel gilt ein Dankeschön für die physikalischen Analysen, bei denen teilweise auch viel persönlicher Einsatz erforderlich war.

Der Abteilung RM von Herrn Dr. A. Preußger danke ich besonders für die finanzielle Unterstützung bei Konferenzteilnahmen. An die Abteilung RM DIE möchte ich ein Dankeschön für die Bereitstellung von Meßkapazität und netzfähigem Computer in der Schlußphase der Arbeit richten.

Herrn Dr. W. Sauert und der Abteilung PI M danke ich für die finanzielle Betreuung in bezug auf Personalkosten, obwohl die Arbeit nicht in derselben Abteilung angesiedelt war.

Dem Bereich Halbleiter der Siemens AG, seit 01.04.1999 Infineon Technologies AG, danke ich für das Promotionsstipendium.

Abschließend möchte ich ganz besonders meinen Eltern für die Unterstützung während des Studiums und der Zeit als Doktorand danken.